

RESEARCH ARTICLE

CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments

Annekatrien Boel, Hanna De Saffel, Wouter Steyaert, Bert Callewaert, Anne De Paepe, Paul J. Coucke and Andy Willaert*

ABSTRACT

Targeted genome editing by CRISPR/Cas9 is extremely well fitted to generate gene disruptions, although precise sequence replacement by CRISPR/Cas9-mediated homology-directed repair (HDR) suffers from low efficiency, impeding its use for high-throughput knock-in disease modeling. In this study, we used next-generation sequencing (NGS) analysis to determine the efficiency and reliability of CRISPR/Cas9-mediated HDR using several types of single-stranded oligodeoxynucleotide (ssODN) repair templates for the introduction of disease-relevant point mutations in the zebrafish genome. Our results suggest that HDR rates are strongly determined by repair-template composition, with the most influential factor being homology-arm length. However, we found that repair using ssODNs does not only lead to precise sequence replacement but also induces integration of repair-template fragments at the Cas9 cut site. We observed that error-free repair occurs at a relatively constant rate of 1-4% when using different repair templates, which was sufficient for transmission of point mutations to the F1 generation. On the other hand, erroneous repair mainly accounts for the variability in repair rate between the different repair templates. To further improve error-free HDR rates, elucidating the mechanism behind this erroneous repair is essential. We show that the error-prone nature of ssODN-mediated repair, believed to act via synthesis-dependent strand annealing (SDSA), is most likely due to DNA synthesis errors. In conclusion, caution is warranted when using ssODNs for the generation of knock-in models or for therapeutic applications. We recommend the application of in-depth NGS analysis to examine both the efficiency and error-free nature of HDR events.

This article has an associated First Person interview with the first author of the paper.

KEY WORDS: CRISPR/Cas9, HDR, Homology-directed repair, Zebrafish, Next-generation sequencing

Center for Medical Genetics, Department of Biomolecular Medicine, Ghent University, Corneel Heymanslaan 10, 9000 Ghent, Belgium.

*Author for correspondence (Andy.Willaert@UGent.be)

© A.B., 0000-0003-3256-7020; B.C., 0000-0002-9743-4205; A.D.P., 0000-0002-4134-5940; P.J.C., 0000-0001-7119-586X; A.W., 0000-0002-9543-1932

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Received 27 April 2018; Accepted 31 August 2018

INTRODUCTION

Site-specific genome-editing technologies enable the efficient generation of knock-out model systems. However, particularly for disease modeling, there is a growing necessity to complement knock-out models with more precise knock-in models, which will accommodate two needs. First, knock-in models are relevant to study diseases caused by specific and/or recurrent point mutations with dominant-negative, hypomorphic or gain-of-function effects. Second, this approach could serve as an *in vivo* tool to assess the pathogenicity of newly identified variants of unknown significance (VUS), which are massively identified by whole-exome sequencing. CRISPR/Cas9 genome editing is considered the most promising and versatile technology to create knock-in disease models (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013; Ran et al., 2013a). The CRISPR/Cas9 system induces a double-stranded DNA break (DSB), carried out by the Cas9 nuclease protein, at a specific target site, recognized by the binding of a single-guide RNA (sgRNA) molecule. Following DSB formation, two main endogenous repair mechanisms can be initiated: either the error-prone non-homologous end joining (NHEJ) pathway, often leading to the introduction of indel (insertion/deletion) mutations, or the homology-directed repair (HDR) pathway. HDR is only activated in the presence of a homologous repair template, naturally provided as the sister chromatid during the G2 and S phase of the cell cycle, leading to the reconstitution of the original sequence. The knock-in modeling procedure exploits this mechanism by supplying the CRISPR/Cas9 system with an artificial repair template, homologous to the target sequence and containing base-pair substitutions or gene insertions of interest. Various types of HDR repair templates have been used so far, including both circular and linear double-stranded DNA (dsDNA) molecules, and single-stranded oligodeoxynucleotides (ssODNs) (Salsman and Dellaire, 2017). For large-scale generation of knock-in models, preference is given to the latter as the design and production of ssODNs is easier, less time-consuming and cheaper than the generation of double-stranded templates such as plasmids (Leonetti et al., 2016; Mikuni et al., 2016). Additionally, recently, methods such as 'easi-CRISPR' have been developed for the generation of long ssODNs consisting of more than 1 kb of sequence, enabling the insertion of longer sequences such as reporters or gene tags (Miura et al., 2018). Moreover, with the use of ssODNs, illegitimate random integration of the exogenous DNA products into the organism's genome is expected to be less frequent as opposed to the use of dsDNA templates (Won and Dawid, 2017; Würtele et al., 2003; Zorin et al., 2005). Finally, single-stranded templates have shown superiority over similarly designed dsDNA templates in terms of HDR efficiency (Beumer et al., 2013; Miura et al., 2015; Yoshimi et al., 2016).

CRISPR/Cas9-induced HDR-mediated knock-in faces two major problems. First, the mechanism can be error prone.

Indeed, imprecise repair events, such as targeted base-pair substitutions combined with insertions and deletions, have been reported (Gratz et al., 2013; Hoshijima et al., 2016; Hwang et al., 2013; Paix et al., 2017; Rivera-Torres et al., 2017; Zu et al., 2013), but the origin and extent of these erroneous repair events has not been thoroughly investigated. Second, its efficiency is relatively low compared to gene disruptions through indel generation. Several cell-endogenous DSB repair mechanisms, including NHEJ and HDR pathways, exist and interact with each other, although the molecular determinants of these interactions are not yet fully known (Kakarougkas and Jeggo, 2014). The decision on which repair pathway is activated is influenced by many factors, including the phase of the cell cycle, DSB complexity, chromatin structure (Brandsma and Gent, 2012; Frank-Vaillant and Marcand, 2002), and the composition and concentration of the repair template (Lin et al., 2014; Rivera-Torres et al., 2017). Previous research explored the impact of adaptations to the length, symmetry and strand complementarity of the ssODN repair template on genome-editing efficiency, although consensus is lacking concerning the impact of these different adaptations (Richardson et al., 2016; Yang et al., 2013). Added chemicals that either block the NHEJ pathway (SCR7, NU7441 and KU0060648) (Ma et al., 2016; Maruyama et al., 2015; Singh et al., 2015; Srivastava et al., 2012) or stimulate the HDR pathway (RS-1, L755507) (Jayathilaka et al., 2008; Pinder et al., 2015; Song et al., 2016; Yu et al., 2015) improve HDR rates in cellular systems to a certain extent, but still need further evaluation in other models.

The zebrafish is frequently used for disease modeling because of its many advantages, such as a rapid development, large offspring numbers, and the ease and speed in generating mutant lines. In this study, we evaluated the efficiency of point-mutation knock-in approaches in zebrafish models. We evaluated the effect of length, symmetry and strand complementarity of ssODN repair templates on HDR-mediated knock-in rates at multiple sgRNA target sites, and assessed a further impact of multiple chemical compounds that either block NHEJ or stimulate HDR. With next-generation sequencing (NGS), we accurately determined precise and erroneous knock-in efficiencies. Our work provides new insights into the effect of repair-template composition on HDR efficiency and stresses the importance of rigorous analysis of CRISPR/Cas9-edited genomic targets to evaluate the error-free nature of the obtained knock-in.

RESULTS

HDR is prone to errors and influenced by ssODN repair-template composition

A prerequisite for activation of the HDR pathway is the presence of a suitable repair template. For each of four different sgRNA target sites in the zebrafish genome (located in the *smad6a*, *tprkb*, *pls3* and *slc2a10* genes), we evaluated multiple ssODN repair-template compositions, altering length, strand complementarity and symmetry, in order to maximize HDR efficiency (Fig. 1A). The results reveal a major impact of repair-template composition on HDR efficiency. In three out of four sgRNA targets, the HDR rate for symmetrical templates increased significantly when extending the template length from 60 to 120 nucleotides (nt), with the maximum improvement being tenfold. Lengthening the template further to 180 nt moderately decreased HDR rates in most cases (Fig. 1B, Table S1). Overall, we obtained maximal total HDR rates ranging from 4 to 8%, depending on the target site. Strand complementarity did not significantly affect HDR rates, although templates that share sequence identity with the strand that is not binding the sgRNA ('non-target' templates) tend to perform slightly

better. Asymmetrically designed templates did not result in better HDR rates compared to symmetrically designed templates: two asymmetrical repair templates, a 'non-target' template with its long homology arm at the protospacer adjacent motif (PAM)-distal side ('NT 120 A left') and a 'target' template with its long homology arm at the PAM-containing side ('T 120 A right'), performed at similar levels compared to the 120 nt symmetrical templates; however, their counterparts with different strand complementarities ('NT 120 A right' and 'T 120 A left') performed remarkably worse and resulted in the lowest repair rates observed.

Further dissection of the nature of the repair events using in-depth NGS data analysis revealed that ssODN-mediated DSB repair does not only lead to scarless genome editing with introduction of the intended base pair change(s), but unexpectedly also causes erroneous integration of ssODN repair-template fragments at the site of the DSB. Consequently, HDR rates dropped from 4–8% (total) to 1–4% (perfect) when removing all NGS reads containing erroneous ssODN integration events from the analysis (Fig. S1). These reads could only be partially aligned to the reference genome. The parts of the reads that were clipped contained fragments of the targeted gene sequence, indicating multiple and often partial erroneous integrations of the repair template. We schematized a number of erroneous integration events of the *smad6a* 'NT 120 S' and 'T 120 S' repair templates to illustrate the wide diversity and complexity of integration patterns (Fig. 2, Fig. S2, Table S2). These patterns ranged from an inverse integration of a part of the repair template to the introduction of multiple repair-template fragments that can be present in either orientation. Additionally, in some cases, a partial deletion of the sequence at the targeted locus was observed.

Remarkably, the observed differences in total HDR rates between the different symmetrical repair templates can be largely attributed to the difference in frequency of erroneous integration events (Fig. 1B, Fig. S1, Table S1). In contrast, the difference in HDR rates between the asymmetrical repair templates did not change when removing all NGS reads containing erroneous ssODN integration events; 'NT 120 A left' and 'T 120 A right' still outperformed their counterparts with different strand complementarities, when only considering perfect repair rates. In addition, for three out of the four sgRNA target sites, perfect repair rates for these 'superior' asymmetrical templates ('NT 120 A left' and 'T 120 A right') are higher than for their symmetrical counterparts ('NT 120 S' and 'T 120 S'). Nevertheless, no specific type of repair template performed consistently best across all four sgRNA target sites.

Inverse relation between distance of base-pair substitution relative to Cas9 cut site and knock-in efficiency

The presence of multiple synonymous nucleotide substitutions in the repair templates (Fig. 3A) allowed us to investigate the impact of the position of the substituted base pair relative to the Cas9 cut site on perfect repair rates (Fig. 3B, Table S3). Successful introduction of the substitution was generally highest at the positions close to the Cas9 cut site, and slightly dropped proportionally with increasing distance. The extent of this drop in HDR rates varied for the different sgRNA target sites. Comparing the incorporation rate of the nucleotide alteration at the position furthest (–15 or –14) and closest (1 or 2) from the Cas9 cut site, averaged for all repair templates of a specific gene, revealed a decrease in perfect repair rates that is twofold for *tprkb*, fourfold for *pls3* and *smad6a*, and 24-fold for *slc2a10*. The biggest reduction in nucleotide alteration rates at the edge positions was noted for some of the repair templates that contain short homology arms ('T 60 S', 'NT 60 S' and the asymmetrical templates).

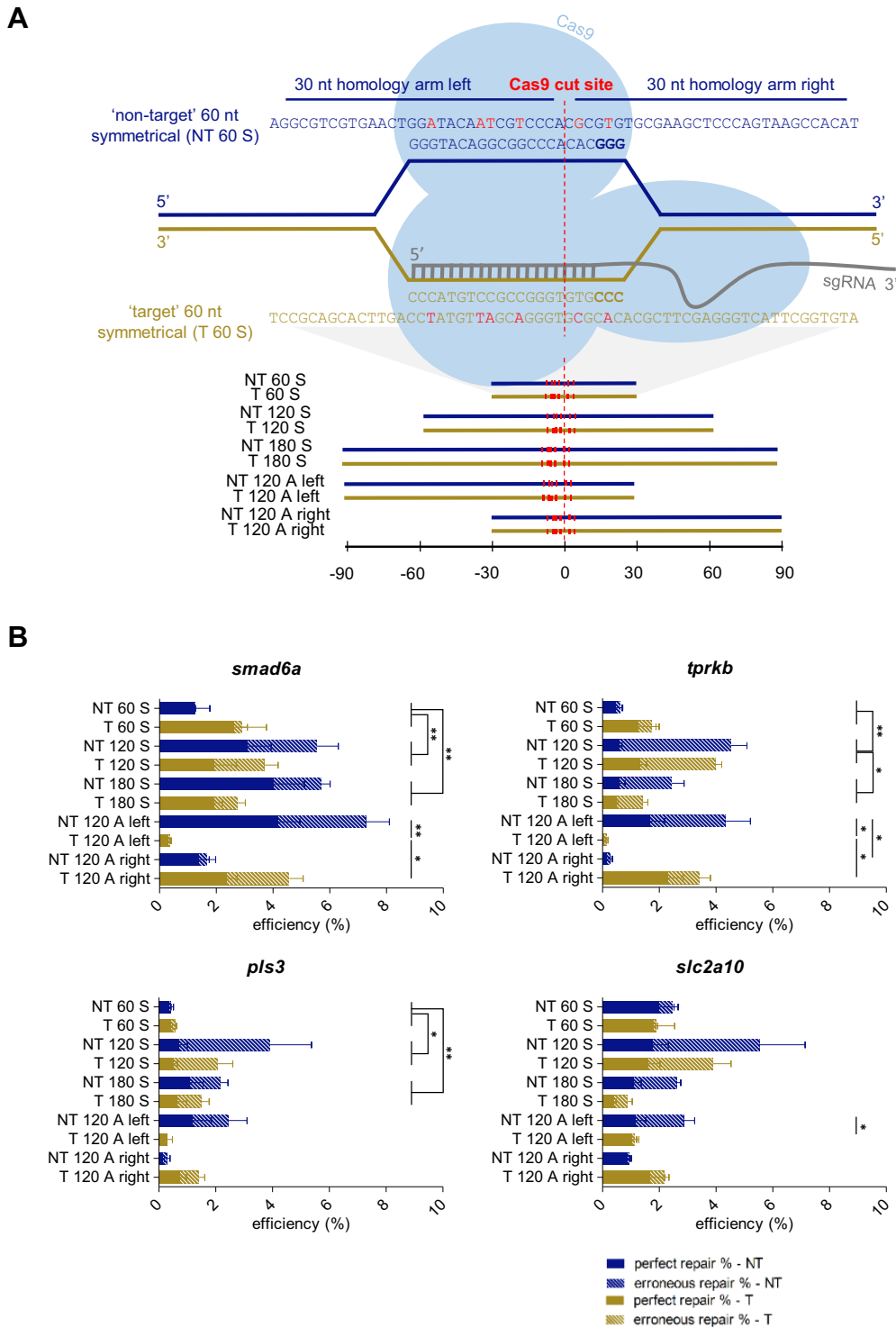


Fig. 1. See next page for legend.

Chemical compound administration does not influence HDR rates in zebrafish

We aimed to increase HDR rates in zebrafish by the use of chemical compounds that, in cell cultures, were shown to either block the NHEJ pathway (SCR7, NU7441, KU0060648) or stimulate the HDR pathway (RS1, L755507). We either co-injected these compounds with Cas9/sgRNA complexes targeting one of the four genes together with a corresponding symmetrical 120 nt repair template ('NT 120 S') or incubated injected embryos for 24 h in

compound screening medium, containing the chemical compounds. For injection, the administered compound dose depended either on compound solubility [selection of the highest compound concentration without aggregate formation in the injection mix (L755507)], compound toxicity [selection of the highest compound concentration without any influence on embryo development and survival (SCR7 and RS1)] and practical considerations [use of the maximal free volume in the injection mix used for compound administration (NU7441 and KU0060648)]. For incubation, the

Fig. 1. Impact of repair-template homology-arm length, strand complementarity and symmetry on HDR efficiency.

(A) Illustration of an sgRNA target site for the zebrafish *smad6a* gene, with protospacer sequence GGGTACAGGCGGCCACAC. Following Cas9 recruitment and sgRNA binding, the DNA is cleaved 3 bp upstream of the PAM sequence (NGG, displayed in bold), which is visualized by the dashed red line and indicated as the 'Cas9 cut site'. We designed ten repair templates, 60, 120 or 180 nucleotides (nt) in length, either corresponding to the sgRNA target strand ('target' – T) or to the complementary strand ('non-target' – NT). The repair templates were either symmetrically (S) or asymmetrically (A) positioned around the Cas9 cut site. The position of the different repair templates relative to the Cas9 cut site is depicted in the lower panel. The sequence of the 60 nt repair templates for *smad6a* is shown in the upper panel, and the nucleotide sequences of all other repair templates are listed in Table S8. Each repair template contains several synonymous nucleotide changes, depicted in red, relative to the reference sequence, including replacement of a guanine nucleotide in the PAM, whenever possible. (B) For each target site (*smad6a*, *tpkbb*, *pbl3* or *slc2a10*) and each repair-template type (NT 60 S, NT 120 S, NT 180 S, NT 120 A left, NT 120 A right, T 60 S, T 120 S, T 180 S, T 120 A left, T 120 A right), average total HDR efficiencies resulting from five independent experiments were plotted. Calculated repair rates represent the number of sequencing reads containing the intended base-pair substitution closest to the Cas9 cut site. The HDR rates were split up into two categories: 'perfect repair %', representing the percentage of NGS reads containing at least the base-pair change closest to the Cas9 cut site (plain bars), and 'erroneous repair %', representing the reads containing erroneous integration events of repair-template fragments (dashed bars). NT, non-target; T, target; 60-120-180, total repair-template length; S, symmetrical; A, asymmetrical. Error bars represent the s.e.m. for five independent biological replicates each consisting of a pooled sample of 20 embryos. Repair rates depicted in this graph are listed in Table S1. Statistical tests performed: one-way ANOVA with blocking (60 nt vs 120 nt, 60 nt vs 180 nt, 120 nt vs 180 nt, target vs non-target) for symmetrical templates, non-parametric Kruskal–Wallis test, followed by pairwise comparison with Dunn–Bonferroni correction, for asymmetrical templates; * $P < 0.05$ and ** $P < 0.01$.

maximal compound concentration for which more than 80% of the embryos developed normally was selected. None of the administered compounds had a notable influence on HDR rates, neither for compound injection (Fig. 4, Table S4) nor for compound incubation (Fig. S3, Table S5).

Precise base-pair substitutions introduced via ssODNs show germline transmission

We investigated the efficacy of germline transmission of base-pair substitutions introduced via HDR-mediated knock-in using ssODN repair templates. At the age of 3 months, we collected germ cells from eight fish per target (*smad6a* and *slc2a10*) and performed NGS. Analysis of the nucleotide alteration located closest to the Cas9 cut site revealed a relevant number of germ cells with HDR events in, respectively, five out of eight and two out of eight screened fish (Fig. 5, Table S6). Most of these founder fish contained germ cells with both perfect repair and erroneous events but, for both gene targets, we identified at least one founder fish with scarless HDR rates above 10%. This provides a good chance of transmission of the nucleotide alteration to the next generation, despite the relatively low perfect repair rates detected during sequence analysis of injected embryos at 1 dpf.

DISCUSSION

The generation of knock-in models has been greatly simplified by the emergence of CRISPR/Cas9-mediated HDR approaches, where genomic target sequences are replaced by homologous donor repair templates. Nevertheless, the efficiency of this process is low, precluding its use for high-throughput disease modeling or therapeutic applications. This study evaluates different approaches

to improve the efficiency of scarless CRISPR/Cas9-mediated knock-in of single-base-pair substitutions by using ssODN repair templates in zebrafish, and accurately assesses total, 'erroneous' and 'perfect' repair rates by deep-sequencing analysis.

In a first step, we evaluated the most effective composition of the ssODN repair template for four different sgRNA target sites in the zebrafish genome. Repair-template length significantly affects total repair rates, with 120 nt being identified as the most optimal length to promote HDR. Shorter templates performed significantly worse, while longer templates showed a moderate drop in HDR rates. These findings are comparable to results obtained in human stem cells (Yang et al., 2013), pigs (Wang et al., 2016) and mice (Shen et al., 2013), where reported optimal total lengths approximate 90, 120 and 100 nt, respectively. We obtained average total HDR-mediated knock-in rates of 4–8% depending on the sgRNA target site when using 120 nt templates. This is more or less comparable to the results of Hruscha et al. (2013), to our knowledge the only other study in zebrafish evaluating HDR using NGS on whole-embryo DNA extracts. Other studies that investigated nuclease-mediated HDR in zebrafish only report on percentages of embryos containing at least one precise editing event, or make estimations on HDR efficiency based on PCR- or restriction-assay-derived data, impeding a direct comparison with our results (Albadri et al., 2017). The inclusion of repair templates that are complementary to either the sgRNA target or non-target strand enabled us to investigate the impact of strand preference on HDR rates. Our data do not support a clear strand preference for symmetrical repair templates at four different sgRNA target sites. This is in agreement with data obtained in other organisms (Lin et al., 2014; Renaud et al., 2016), but does not rule out specific locus-dependent strand preference as suggested previously (Hwang et al., 2013).

Next, we examined the impact of repair-template symmetry on HDR rates. A recent study has suggested that use of asymmetrical templates leads to improved HDR rates; Richardson et al. (2016) showed that, in cell cultures, complete dissociation of the Cas9-DNA complex after DSB creation is preceded by an early release of the PAM-distal non-target strand (Richardson et al., 2016). This insight might contribute to rational repair-template design, particularly by constructing asymmetrical templates complementary to the non-target strand, and with homology arms of 90 nt and 30 nt, corresponding respectively to the PAM-containing and the PAM-distal side of the DSB. Exploration of the performance of this specific repair template in zebrafish, alongside a number of other asymmetrical templates complementary to either the 'target' or 'non-target' strand, did not confirm that the abovementioned optimal asymmetrical repair template (designated 'T 120 A right' in our study), or any other asymmetrical template, performed notably better than the symmetrical repair templates. Interestingly, though, significant differences within the set of asymmetrical repair templates were noticeable, with 'NT 120 A left' and 'T 120 A right' clearly outperforming the other two asymmetrical repair templates. This remarkable difference, together with our finding that short 60 nt templates perform significantly worse than 120 nt or 180 nt templates, adds to our understanding of the mechanisms involved in DSB repair. A currently emerging hypothesis suggests that DSB repair using single-stranded templates proceeds through the single-stranded template repair (SSTR) pathway (Jasin and Haber, 2016). This process occurs independently of RAD51 and BRCA2, the key components of homologous recombination (HR) (Bothmer et al., 2017). It is proposed that the mechanism behind SSTR is based on synthesis-dependent strand annealing (SDSA), consisting of three

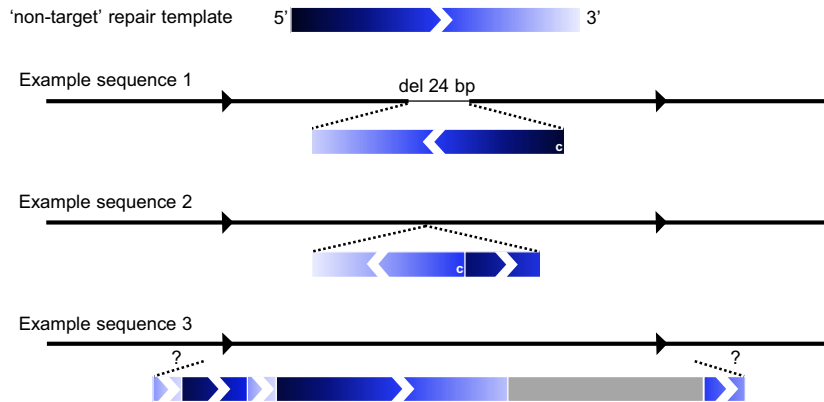
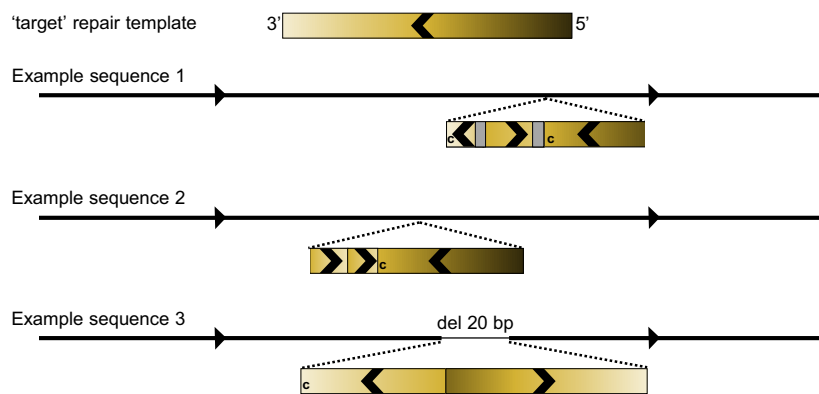
NT 120 S**T 120 S**

Fig. 2. Simplified schematic representation of erroneous repair-template integration events at the *smad6a* zebrafish gene. For both the 'NT 120 S' and 'T 120 S' repair templates, three examples of NGS reads are schematized to clarify the erroneous repair patterns that were encountered in this study. For each example, the black line with arrows represents the reference sequence, and the two black dashed lines indicate the approximate location of repair-template insertions. Inserted fragments are depicted as rectangles, colored in blue or yellow, in accordance with the used repair template (indicated in the scheme as 'non-target' or 'target' repair template on top of the three example sequences). Gray rectangles depict 'random' sequences not corresponding to the repair-template sequence. The color gradient clarifies to which part of the repair template the insertional fragment corresponds and defines the orientation of the fragment. This is additionally illustrated by arrows overlaying the rectangles. A 'c' in the rectangle means that it corresponds with the repair template's complementary sequence. In 'NT 120 S' example sequence 3, the integrated repair template fragments could not be assigned to a certain location in the reference sequence, due to the limited length of NGS sequence reads (250 bp), which is depicted by two question marks ("?"). This scheme is a simplified representation. A more detailed scheme of these examples, using the NGS read sequence as reference point, is depicted in Fig. S2.

major steps: (1) resection of the DSB to create 3' overhangs, (2) pairing of the overhangs with the donor DNA followed by extension of the 3' strand by DNA-polymerase-mediated DNA synthesis (initiation), and (3) strand displacement of the extended 3' strand from the template strand with reannealing of the extended 3' strand and the endogenous DNA at the opposite end of the DSB (resolution) (Fig. 6A). This process is characterized by unidirectional sequence conversion (Davis and Maizels, 2016; Kan et al., 2017; Paix et al., 2017). Templates containing a short (30 nt) homology arm at the side of the repair template that does not anneal to one of 3' overhangs ('NT 60 S', 'T 60 S', 'T 120 A left', 'NT 120 A right'), but participates in the resolution step, perform poorly in our study (Fig. 6B). These findings point towards a necessity for these particular homology arms to have an optimal length (about 60 nt) for resolution, at least in the zebrafish. The difference in performance between the asymmetrical templates included in this study may thus be linked to SDSA pathway requirements, instead of being attributed to the nature of the dissociation of the Cas9-DNA complex after DSB, as suggested by Richardson et al. (2016).

The outcome of precise genome editing using ssODNs depends on the perfect replacement of the target sequence by the repair template. Extended data analysis of the obtained deep-sequencing results in this study revealed that approximately half of the sequencing reads containing the base-pair alterations of interest additionally show a complex pattern of multiple integrated fragments of the repair template that can be present in both directions. This finding was not entirely unexpected, since a number of studies have reported imprecise HDR edits (Gratz et al., 2013;

Hwang et al., 2013; Paix et al., 2017; Rivera-Torres et al., 2017; Zu et al., 2013), but the extent of this phenomenon was unanticipated. It is generally believed that the use of ssODNs, unlike dsDNA templates, circumvent random and illegitimate integration of the exogenous DNA products into the organism's genome (Won and Dawid, 2017; Würtele et al., 2003; Zorin et al., 2005). Therefore, the erroneous integration events encountered in this study are most likely attributed to mechanisms other than random integration. Since ssODN-induced DSB repair most likely occurs through SDSA (Paix et al., 2017), initiating DNA replication steps may be prone to polymerase slippage and stalling (Rodgers and McVey, 2016) with subsequent dissociation of unstable intermediates and template switching (Rodgers and McVey, 2016). Multiple rounds of template switching can lead to complex mutational spectra (Guirouilh-Barbat et al., 2014). Correspondingly, template switching during CRISPR/Cas9-mediated HDR in HEK293T cells has been recently demonstrated by Paix et al. (2017). We hypothesized that the presence of microhomologies in the target sequence may explain why certain targets are prone to template switching with subsequent erroneous repair, as previously seen in genomic rearrangement events (Löytynoja and Goldman, 2017) (Fig. 6C). As a proof-of-principle, we showed, for one of the repair templates targeting the *smad6* gene, that erroneous repair occurred through multiple template switching events during SDSA-based DSB repair, based on the presence of short microhomology sequences (Fig. S4). The sequence-specific nature of these events holds promise for future development of prediction algorithms for erroneous repair in CRISPR/Cas9-mediated HDR.

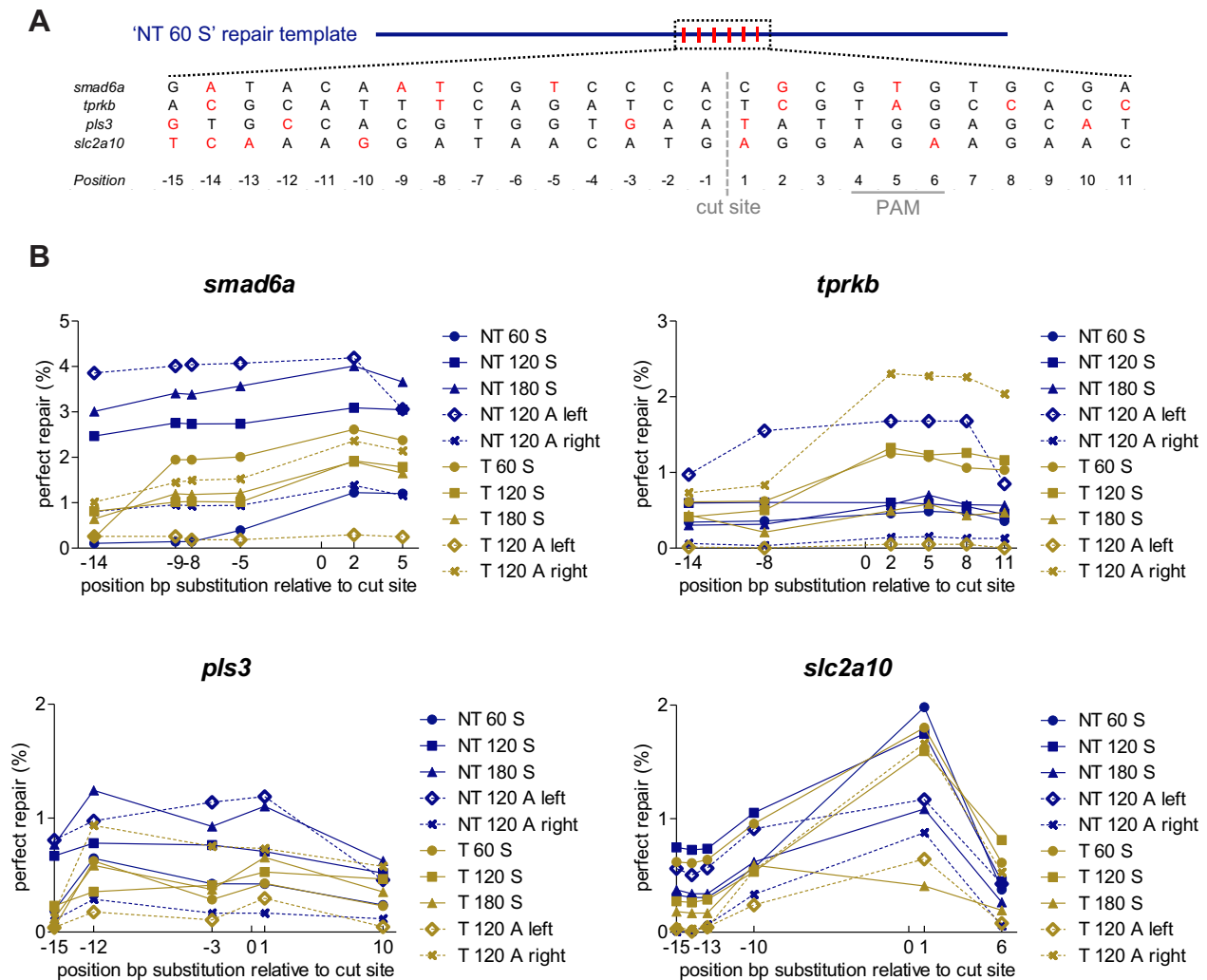


Fig. 3. Impact of the position of the substituted nucleotide in the repair template on repair rates. (A) Repair-template composition with specified locations of the base pair alterations, shown for an 'NT 60 S' repair template. Each repair template included in this study contains several synonymous nucleotide changes relative to the reference sequence. Whenever possible, a guanine nucleotide of the PAM was replaced, in order to avoid undesired cleavage of the repair template by Cas9. In addition, five nucleotide changes were included in the region surrounding the Cas9 cut site (designated as point zero, visualized with a gray dashed line). All nucleotide changes are depicted in red. The position of the PAM sequence is marked with a full gray line. (B) The HDR efficiency (perfect repair only) for each base-pair substitution in each repair template is plotted as a function of the relative distance to the Cas9 cut site. Each plotted data point shows average values of five independent experiments, each analyzing 20 embryos. To improve clarity, error bars were omitted. NT, non-target; T, target; 60-120-180, total repair-template length; S, symmetrical; A, asymmetrical. Repair rates depicted in this graph are listed in Table S3.

Importantly, no repair template performs unambiguously best across all sgRNA target sites, when solely considering perfect repair rates, although there is a tendency for some asymmetrical repair templates ('NT 120 A left' and 'T 120 A right') to outperform their symmetrical counterparts ('NT 120 S' and 'T 120 S'), comparable to previously reported results in zebrafish (Prykhodzhiy et al., 2018). Therefore, it is advisable to try different types of repair templates when attempting to knock-in a point mutation of interest.

Our data clearly indicate that the outperformance of some of the templates when considering total HDR rates was mainly due to erroneous repair. Excluding the sequencing reads that contain erroneous integration events leads to a drop in HDR rates (of the base pair located closest to the DSB site) from 4–8% to 1–4%, depending on the sgRNA target site. Moreover, these perfect repair rates drop even further if the distance between the base alteration and the Cas9 cut site increases, a phenomenon that was observed previously (Elliott et al., 1998; Paquet et al., 2016; Ran et al.,

2013b). This implies that the distance between the Cas9-induced DSB and the intended sequence alteration should be minimized. Despite the low rates of scarless repair in zebrafish embryos, efficient transmission of nucleotide alterations to subsequent generations was found to be relatively efficient.

Finally, we evaluated HDR rates following chemical intervention blocking the NHEJ pathway (SCR7, NU7441, KU0060648) or stimulating the HDR pathway (RS1, L755507), an approach reportedly successful in cellular systems. Two delivery methods (incubation and injection) were applied. Surprisingly, none of these compounds affected either perfect or erroneous HDR rates in zebrafish. Possible explanations include: (1) delivery through injection or compound toxicity in the zebrafish hampers the delivery of the minimal compound concentration necessary for its intended activity, (2) the absence of a response to compound administration could be attributed to species differences, although it is known that DSB repair mechanisms are generally well conserved,

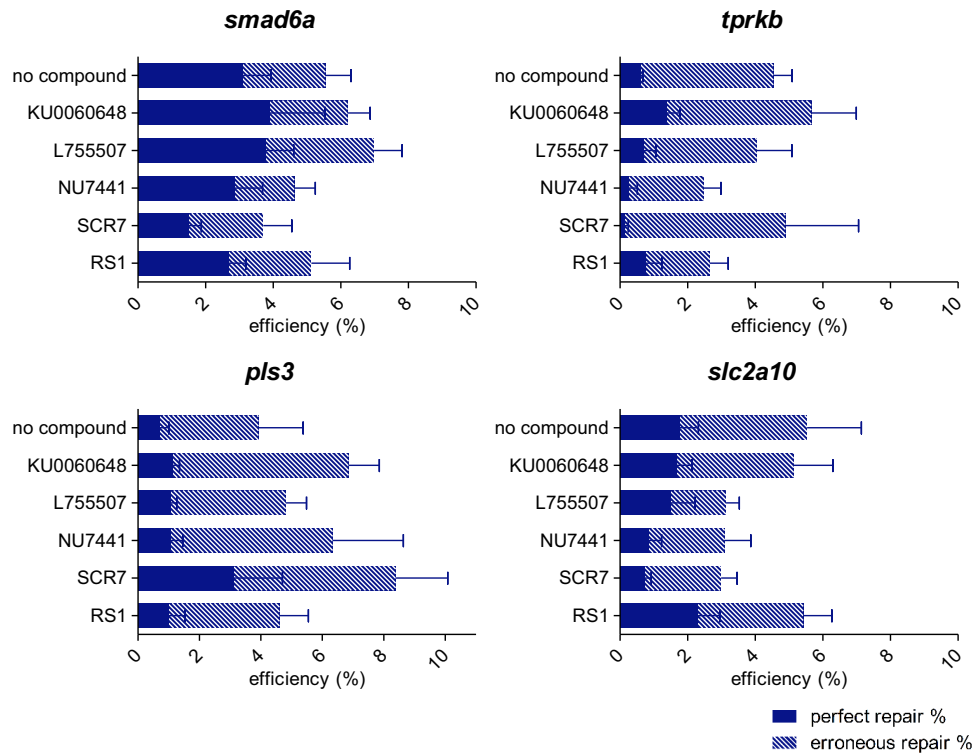


Fig. 4. Influence of chemical compound administration through injection on HDR efficiency. Injection mixes containing the NT 120 S repair template were complemented with chemical compounds that either inhibit specific components of the NHEJ pathway, including SCR7, NU7441 and KU0060648, or that were shown to stimulate the HDR pathway, including RS1 and L755507. Five independent experiments were carried out and average total HDR rates are shown, split into two categories: perfect repair % (plain bars) and erroneous repair % (dashed bars). Error bars represent the s.e.m. for five independent biological replicates each consisting of a pooled sample of 20 embryos. Repair rates depicted in this graph are listed in Table S4. Statistical tests performed: independent samples *t*-test for normal distributed groups and the non-parametric independent samples Mann–Whitney *U*-test for non-normal distributed groups.

or (3) the RAD51 stimulator RS1 might only have an effect when using double-stranded repair templates, since single-stranded template repair of DSBs is believed to act independently from RAD51 (Bothmer et al., 2017). In view of the latter, compounds that induce the SSTR pathway could have better potential in increasing HDR rates when using single-stranded templates. Interestingly, a recent study in zebrafish has shown that SCR7 administration through incubation increased HDR rates from 5 to 13% using an ssODN donor (Zhang et al., 2018). One possible reason that our results conflict with their conclusion is the difference in analysis method. We report repair rates that match the actual frequency of altered alleles in a pool of embryos, determined by NGS, while the study of Zhang and co-workers only reports the percentage of embryos containing HDR events. Moreover, these HDR rates are determined by a method solely based on PCR amplification. The use of PCR-based techniques using primers that specifically recognize the mutant sequence introduced by HDR should be avoided since they cannot distinguish erroneous repair from perfect repair and may thus lead to an overestimation of error-free HDR events. In addition, this approach lacks sensitivity to detect small differences in HDR rates. Another study performed in *Xenopus* showed that, similarly to our study, SCR7 was toxic to embryos and did not improve HDR rates (Aslan et al., 2017). Administration to non-dividing oocytes, on the other hand, appeared to significantly increase HDR rates. Therefore, the procedure of oocyte treatment could be a challenging but promising approach for future experiments in zebrafish, aiming to increase HDR rates (Xie et al., 2016).

Overall, in this study we tested the performance of multiple strategies in the zebrafish model system that were previously shown to improve HDR rates in cellular systems, including optimal donor template design and chemical compound administration. We demonstrate that HDR-mediated knock-in is influenced by the design of the ssODN template, while none of the investigated chemical compounds significantly increased HDR rates in zebrafish. Importantly, we identified that imprecise complex genomic insertions of the donor template constitute a significant portion of total HDR events, and suggest that repeated template switching during SDSA is the underlying mechanism of erroneous ssODN-mediated HDR. These findings stress the need for optimal experimental design and reliable assessment of HDR-mediated knock-in in zebrafish (Box 1).

MATERIALS AND METHODS

sgRNA design and production

We designed sgRNA molecules targeting four different zebrafish genes (*smad6a*, *tprkb*, *pls3* and *slc2a10*) with CRISPRdirect software (<http://crispr.dbcls.jp/>) (Naito et al., 2015). For *in vitro* transcription, we synthesized gBlocks with the following structure (IDT, gBlock): 5'-CCG-CTAGCTAATACGACTCACTATA-**GG-N₁₈**-GTTTATAGAGCTAGAAA-TAGCAAGTTAAATAAGGCTAGTCCGTTATCACTTGAAAAAG-TGGCACCAGATCGGTGCTTTT-3', where the two guanines in bold fulfill the T7 polymerase promoter requirement and N₁₈ represents the last 18 nucleotides of the protospacer sequence. Specific gBlock sequences for the four targets are listed in Table S7. Lyophilized gBlock molecules were dissolved in nuclease-free water at a concentration of 10 ng/μl and 4 μl was used as an input for *in vitro* transcription, using the MEGAshortscript™ T7 Transcription Kit

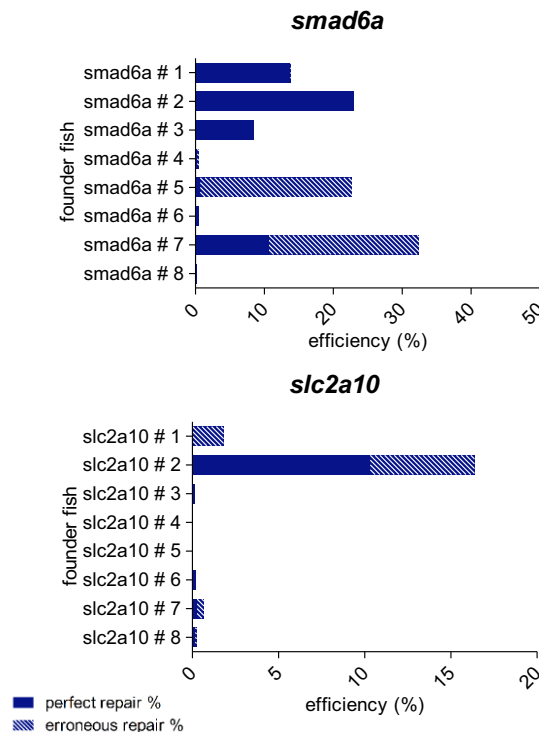


Fig. 5. Determination of germline transmission of precise base-pair substitutions introduced by CRISPR/Cas9-mediated HDR. For two of the four target sites included in this study (the *slc2a10* and *smad6a* zebrafish genes), embryos injected with injection mixes containing the NT 120 S repair template were grown until adulthood. For each target site, eight adult fish (labeled founder fish # 1, # 2, ... # 8) were screened for the presence of the precise base-pair substitutions located closest to the Cas9 cut site in their germ cells. Therefore, DNA was extracted from collected eggs or sperm, and subjected to NGS analysis. HDR rates are shown, split into two categories: perfect repair % (plain bars) and erroneous repair % (dashed bars). Repair rates depicted in this graph are listed in Table S6.

(Invitrogen, cat. no. AM1354). We followed the general guidelines of the manufacturer except for the duration of the incubation step at 37°C, which was carried out overnight to obtain a maximum yield. The transcription reaction was purified using the MEGAclean™ Kit (Life Technologies, cat. no. AM1908), following the manufacturer's instructions. Quantity and integrity of the RNA was determined using, respectively, the DropSense96 device (Trinean) and the Experion microfluidic capillary electrophoresis system (Bio-Rad). The RNA was aliquoted and stored at -80°C.

Repair-template design

For each of the four different sgRNA target sites (located in the *smad6a*, *tpk3b*, *pls3* and *slc2a10* genes), we designed ten ssODN repair templates (Fig. 1A). Six templates contained symmetrical homology arms, flanking the theoretical Cas9 cut site (located 3 base pairs upstream of the PAM), with total lengths of 60 nt, 120 nt or 180 nt and sharing sequence identity either with the sgRNA-binding strand (often referred to as 'antisense', here designated as 'target': 'T 60 S', 'T 120 S', 'T 180 S'), or with the complementary strand (often referred to as 'sense', here designated as 'non-target': 'NT 60 S', 'T 120 S', 'T 180 S'). Four asymmetrical repair templates consisted of 30 nt and 90 nt homology arms, corresponding, respectively, to the PAM-distal and PAM-proximal side relative to the theoretical Cas9 cut site or *vice versa* and sharing sequence identity with either the sgRNA target ('T 120 A right' and 'T 120 A left') or non-target ('NT 120 A right' and 'NT A 120 left') strand. Whenever possible, we masked the sequence corresponding to the PAM by replacing a guanine with another nucleotide, introducing a synonymous mutation. Each of the ten repair templates further contained five nucleotide substitutions in a region from 15 nt upstream to 11 nt downstream of the Cas9 cut site. The substitutions

introduce synonymous mutations as well, in order to avoid a possible impact on the gene product and, hence, embryo survival and experimental outcome. The IDT CodonOpt webtool (<https://eu.idtdna.com/CodonOpt>) was used to determine the most optimal synonymous nucleotide changes for zebrafish. These substitutions avoid possible Cas9 cleavage of the newly edited endogenous DNA sequence, and enable correlation analysis between the mutation-to-DSB distance and HDR efficiency. The sequence and a schematic overview of all ssODN molecules are displayed in Table S8 and Fig. 1A. The ssODNs were ordered as ultramer oligonucleotides, without PAGE purification (IDT, 4 nmol) and were dissolved at a concentration of 10 µM.

Danio rerio (zebrafish) maintenance and injection procedures

We adhered to the general guidelines, in agreement with EU Directive 2010/63/EU for animals, for zebrafish handling, mating, embryo collection and maintenance (Lawrence, 2007; Westerfield, 2000). Approval for this study was provided by the local committee on the Ethics of Animal Experiments (Ghent University Hospital, Ghent, Belgium; Permit Number: ECD 14/31 and ECD 17/41). Following fertilization, we micro-injected one-cell-stage AB zebrafish embryos in the cell with 1.4 nl injection mix composed of 25 pg sgRNA, 250 pg Cas9 protein (ToolGen, Cas9 wild-type nuclease protein with NLS), 2 µM ssODN template, RNase-free water and Phenol Red sodium salt indicator (Sigma Aldrich, cat. no. P4758). In specific experiments, we complemented this mixture with a chemical compound (see section 'Compound administration' and Table S9).

Compound administration

The following chemical compounds were dissolved in DMSO and either supplemented to the injection mix (Table S9) or to the compound screening medium (Table S10): SCR7 (Xcessbio Biosciences Inc., cat. no. M60082-2s), KU0060648 (APEXbio, cat. no. A1769), NU7441 (Santa Cruz Biotechnology, cat. no. sc-208107), L755507 (Santa Cruz Biotechnology, cat. no. sc-204045), RS1 (Santa Cruz Biotechnology, cat. no. sc-222240). Compound screening medium consisted of 1x E3 medium supplemented with 1 mM tris/HCl (pH 7.4), 0.05 U/ml penicillin and 50 ng/ml streptomycin (Murphy and Zon, 2006). Screening medium was distributed in 96-well plates, each well containing 100 µl medium and one injected embryo. Treatment was performed for 24 h.

Genomic DNA extraction

At 1 day post-fertilization (dpf) we categorized injected embryos as 'dead', 'malformed' or 'normally developing'. We performed genomic DNA extraction on a pool of 20 normally developing embryos. Additionally, normally developing embryos, co-injected with CRISPR/Cas9 components targeting either *slc2a10* or *smad6a* and the corresponding symmetrical 120 nt repair template ('NT 120 S'), were raised to adulthood. Eggs and sperm cells were collected by squeezing female and male fish at the age of 3 months according to the guidelines in 'The Zebrafish Book' (Westerfield, 2000). DNA extraction was carried out using the KAPA Express Extract DNA Extraction Kit (Kapa Biosystems, KK7103). Embryos, eggs or sperm were incubated at 60°C for 10 min and 95°C for 5 min in an extraction mix of 5 µl 10x Kapa Express Extract Buffer, 1 µl Express Extract Enzyme (1 U/µl) and 44 µl PCR grade water. The resulting DNA was stored at -20°C for subsequent PCR amplification.

Sequence analysis

Using target-specific primers, the extracted DNA was singleplex PCR-amplified (Table S11). The resulting PCR products were subjected to the Nextera XT library preparation protocol (Illumina, San Diego, CA) and deep sequenced on a MiSeq instrument (Illumina) using 2x250 bp cycles (De Leeneer et al., 2015). Data analysis was performed using BATCH-GE, a Perl-based bioinformatics tool specifically designed for the assessment of NGS data resulting from target-specific genome-editing experiments (Boel et al., 2016). Adaptations to the original BATCH-GE script were made in order to be able to discriminate between perfect repair and erroneous events. The adapted analysis workflow is schematized in Fig. S5 and input

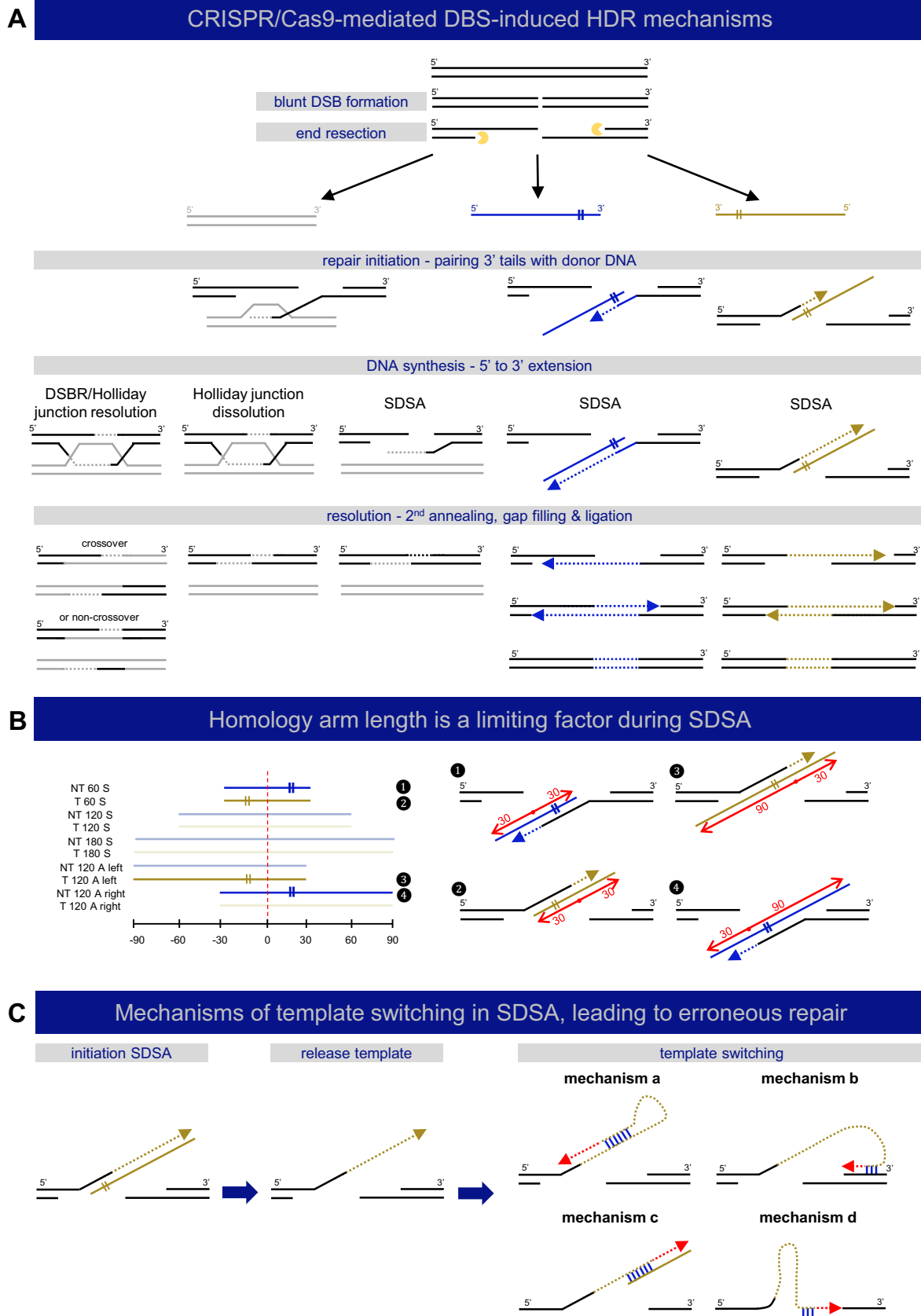


Fig. 6. See next page for legend.

Fig. 6. Mechanisms of DNA double-strand break repair. (A) A schematic representation of different described mechanisms of DSB-induced HDR is shown, categorized by repair template type: double-stranded (gray) or single-stranded ('target' – yellow, 'non-target' – blue). DSB repair by HDR is always initiated by DNA end resection, resulting in 3' single-stranded DNA tails, which participate in strand invasion into homologous sequences. DSB repair with double-stranded repair templates can be carried out by three different mechanisms: double-strand break repair (DSBR or Holliday junction resolution), Holliday junction dissolution, or synthesis-dependent strand annealing (SDSA) (Allers and Lichten, 2001; Szostak et al., 1983). DSB repair using single-stranded templates, sometimes referred to as SSTR, acts via SDSA (Davis and Maizels, 2016; Kan et al., 2017; Paix et al., 2017). (B) Homology-arm length influences the efficiency of SDSA. The left panel depicts the composition of all included repair templates in this study, where the four templates that perform significantly worse in comparison to the others in terms of total repair rates (see Fig. 1B) are highlighted (1–4). The location of the DSB is marked with a red dashed line. These four templates all contain a short (30 nt) homology arm at the side of the repair template that does not anneal to one of 3' overhangs (marked with 'I' sign), but participates in the resolution step, as also depicted in the right panel, where each of these inefficient templates is fitted to the SDSA model for DSB repair with single-stranded templates. The red dots refer to the DSB location, with left and right homology arms marked with red arrows and length indications. (C) Mechanisms of template switching in SDSA leading to erroneous repair. Following SDSA initiation, dissociation of unstable intermediates can lead to template switching. Four possible mechanisms for template switching during SDSA are shown. These mechanisms are based on microhomology (binding marked in blue) and can result in erroneous replication (red dashed arrows).

parameters are shown in Table S12. The adapted script is freely available for academic use and can be downloaded from <https://github.com/WouterSteyaert/BATCH-GE.git>.

Statistical analysis

Statistical analysis was carried out using SPSS statistics 24. We used the following tests: for symmetrical templates: one-way ANOVA with blocking (60 nt vs 120 nt, 60 nt vs 180 nt, 120 nt vs 180 nt, target vs non-target). For asymmetrical templates: non-parametric Kruskal–Wallis test, followed by pairwise comparison with Dunn–Bonferroni correction. For compound testing: independent samples *t*-test for normal distributed groups and the non-parametric independent samples Mann–Whitney *U*-test for non-normal distributed groups. Error bars represent standard error of mean (s.e.m.). We

repeated each experiment five times to level out variability related to the injection procedure.

Competing interests

The authors declare no competing interests.

Author contributions

Conceptualization: A.B., A.W.; Methodology: A.B., H.D.S., A.W.; Software: W.S.; Validation: A.B., H.D.S.; Formal analysis: A.B., H.D.S.; Investigation: A.B., H.D.S.; Data curation: A.B., A.W.; Writing – original draft: A.B., W.S., A.W.; Writing – review & editing: A.B., H.D.S., W.S., B.C., A.D.P., P.J.C., A.W.; Visualization: A.B.; Supervision: B.C., A.D.P., P.J.C., A.W.; Project administration: A.W.; Funding acquisition: B.C., A.D.P., P.J.C., A.W.

Funding

This work was supported by the Ghent University (Universiteit Gent) Methusalem grant BOF08/01M01108 and BOFMET2015000401 to A.D.P. and by Research Foundation – Flanders (Fonds Wetenschappelijk Onderzoek) FWOOPR2013025301. B.C. is a Senior Clinical Investigator of the Research Foundation – Flanders.

Supplementary information

Supplementary information available online at <http://dmm.biologists.org/lookup/doi/10.1242/dmm.035352.supplemental>

References

- Albadri, S., Del Bene, F. and Revenu, C. (2017). Genome editing using CRISPR/Cas9-based knock-in approaches in zebrafish. *Methods* **121–122**, 77–85.
- Allers, T. and Lichten, M. (2001). Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* **106**, 47–57.
- Aslan, Y., Tadjuidje, E., Zorn, A. M. and Cha, S.-W. (2017). High-efficiency non-mosaic CRISPR-mediated knock-in and indel mutation in F0 *Xenopus*. *Development* **144**, 2852–2858.
- Beumer, K. J., Trautman, J. K., Mukherjee, K. and Carroll, D. (2013). Donor DNA utilization during gene targeting with zinc-finger nucleases. *G3* **3**, 657–664.
- Boel, A., Steyaert, W., De Rocker, N., Menten, B., Callewaert, B., De Paepe, A., Coucke, P. and Willaert, A. (2016). BATCH-GE: batch analysis of next-generation sequencing data for genome editing assessment. *Sci. Rep.* **6**, 30330.
- Bothmer, A., Phadke, T., Barrera, L. A., Margulies, C. M., Lee, C. S., Buquicchio, F., Moss, S., Abdulkarim, H. S., Selleck, W., Jayaram, H. et al. (2017). Characterization of the interplay between DNA repair and CRISPR/Cas9-induced DNA lesions at an endogenous locus. *Nat. Commun.* **8**, 13905.
- Brandsma, I. and Gent, D. C. (2012). Pathway choice in DNA double strand break repair: observations of a balancing act. *Genome Integr.* **3**, 9.
- Cong, L., Ran, F. A., Cox, D., Lin, S. L., Barretto, R., Habib, N., Hsu, P. D., Wu, X. B., Jiang, W. Y., Marraffini, L. A. et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823.
- Davis, L. and Maizels, N. (2016). Two distinct pathways support gene correction by single-stranded donors at DNA nicks. *Cell Rep.* **17**, 1872–1881.
- De Leeneer, K., Hellemaers, J., Steyaert, W., Lefever, S., Vereecke, I., Debals, E., Crombez, B., Baetens, M., Van Heetvelde, M., Coppieters, F. et al. (2015). Flexible, scalable, and efficient targeted resequencing on a benchtop sequencer for variant detection in clinical practice. *Hum. Mutat.* **36**, 379–387.
- Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J. A. and Jasin, M. (1998). Gene conversion tracts from double-strand break repair in mammalian cells. *Mol. Cell. Biol.* **18**, 93–101.
- Frank-Vaillant, M. and Marcand, S. (2002). Transient stability of DNA ends allows nonhomologous end joining to precede homologous recombination. *Mol. Cell* **10**, 1189–1199.
- Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., Wildonger, J. and O'Connor-Giles, K. M. (2013). Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**, 1029–1035.
- Guirouilh-Barbat, J., Lambert, S., Bertrand, P. and Lopez, B. S. (2014). Is homologous recombination really an error-free process? *Front. Genet.* **5**, 175.
- Hoshijima, K., Jurynek, M. J. and Grunwald, D. J. (2016). Precise editing of the zebrafish genome made simple and efficient. *Dev. Cell* **36**, 654–667.
- Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C. and Schmid, B. (2013). Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* **140**, 4982–4987.
- Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Kaini, P., Sander, J. D., Joong, J. K., Peterson, R. T. and Yeh, J.-R. J. (2013). Heritable and precise zebrafish genome editing using a CRISPR-Cas system. *PLoS ONE* **8**, e68708.
- Jasin, M. and Haber, J. E. (2016). The democratization of gene editing: insights from site-specific cleavage and double-strand break repair. *DNA Repair* **44**, 6–16.

Box 1. Conclusions and guidelines for design and analysis of CRISPR/Cas9-mediated HDR experiments in zebrafish

Experimental setup

- None of the repair templates included in this study performs unambiguously best in terms of scarless HDR rates (perfect repair).
- Repair-template design determines the extent of erroneous integration of repair-template fragments (erroneous repair).
- Repair templates with different strand complementarities ('target' vs 'non-target') perform similarly.
- Specific types of 120 nt asymmetrical repair templates (designated 'NT 120 A left' and 'T 120 A right' in this study) have a tendency to outperform 120 nt symmetrical repair templates.
- The distance between the Cas9 cut site and the base-pair substitution should be minimized.
- Chemical compound administration (SCR7, NU7441, KU0060648, RS1, L755507), both through injection or by incubation, does not increase HDR rates in zebrafish embryos

HDR analysis

- Methods solely based on PCR amplification (for example with one primer targeting the modified sequence) or restriction digests are not suitable to detect erroneous HDR.
- Next-generation sequencing (NGS) or Sanger sequencing of a region substantially larger than the repair-template length is advised to be able to detect erroneous repair.

- Jayathilaka, K., Sheridan, S. D., Bold, T. D., Bochenska, K., Logan, H. L., Weichselbaum, R. R., Bishop, D. K. and Connell, P. P. (2008). A chemical compound that stimulates the human homologous recombination protein RAD51. *Proc. Natl. Acad. Sci. USA* **105**, 15848-15853.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821.
- Kakarougkas, A. and Jeggo, P. A. (2014). DNA DSB repair pathway choice: an orchestrated handover mechanism. *Br. J. Radiol.* **87**, 20130685.
- Kan, Y., Ruis, B., Takasugi, T. and Hendrickson, E. A. (2017). Mechanisms of precise genome editing using oligonucleotide donors. *Genome Res.* **27**, 1099-1111.
- Lawrence, C. (2007). The husbandry of zebrafish (*Danio rerio*): a review. *Aquaculture* **269**, 1-20.
- Leonetti, M. D., Sekine, S., Kamiyama, D., Weissman, J. S. and Huang, B. (2016). A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc. Natl. Acad. Sci. USA* **113**, E3501-E3508.
- Lin, S., Staahl, B. T., Alla, R. K. and Doudna, J. A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* **3**, e04766.
- Löytynoja, A. and Goldman, N. (2017). Short template switch events explain mutation clusters in the human genome. *Genome Res.* **27**, 1039-1049.
- Ma, Y., Chen, W., Zhang, X., Yu, L., Dong, W., Pan, S., Gao, S., Huang, X. and Zhang, L. (2016). Increasing the efficiency of CRISPR/Cas9-mediated precise genome editing in rats by inhibiting NHEJ and using Cas9 protein. *RNA Biol.* **13**, 605-612.
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. and Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826.
- Maruyama, T., Dougan, S. K., Truttmann, M. C., Bilate, A. M., Ingram, J. R. and Ploegh, H. L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538-542.
- Mikuni, T., Nishiyama, J., Sun, Y., Kamasawa, N. and Yasuda, R. (2016). High-throughput, high-resolution mapping of protein localization in mammalian brain by in vivo genome editing. *Cell* **165**, 1803-1817.
- Miura, H., Gurumurthy, C. B., Sato, T., Sato, M. and Ohtsuka, M. (2015). CRISPR/Cas9-based generation of knockdown mice by intronic insertion of artificial microRNA using longer single-stranded DNA. *Sci. Rep.* **5**, 12799.
- Miura, H., Quadros, R. M., Gurumurthy, C. B. and Ohtsuka, M. (2018). Easy-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors. *Nat. Protoc.* **13**, 195-215.
- Murphy, R. D. and Zon, L. I. (2006). Small molecule screening in the zebrafish. *Methods* **39**, 255-261.
- Naito, Y., Hino, K., Bono, H. and Ui-Tei, K. (2015). CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* **31**, 1120-1123.
- Paix, A., Folkmann, A., Goldman, D. H., Kulaga, H., Grzelak, M. J., Rasoloson, D., Paidemarry, S., Green, R., Reed, R. R. and Seydoux, G. (2017). Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc. Natl. Acad. Sci. USA* **114**, E10745-E10754.
- Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K. M., Gregg, A., Noggle, S. and Tessier-Lavigne, M. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**, 125-129.
- Pinder, J., Salsman, J. and Deltre, G. (2015). Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res.* **43**, 9379-9392.
- Prykhodzhiy, S. V., Fuller, C., Steele, S. L., Veinotte, C. J., Razaghi, B., Robitaille, J. M., McMaster, C. R., Shlien, A., Malkin, D. and Berman, J. N. (2018). Optimized knock-in of point mutations in zebrafish using CRISPR/Cas9. *Nucleic Acids Res.* **46**, e102.
- Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y. et al. (2013a). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389.
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. and Zhang, F. (2013b). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281-2308.
- Renaud, J.-B., Boix, C., Charpentier, M., De Cian, A., Cochenne, J., Duvernois-Berthet, E., Perrouault, L., Tesson, L., Edouard, J., Thinnard, R. et al. (2016). Improved genome editing efficiency and flexibility using modified oligonucleotides with TALEN and CRISPR-Cas9 nucleases. *Cell Rep.* **14**, 2263-2272.
- Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. and Corn, J. E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339-344.
- Rivera-Torres, N., Banas, K., Bialk, P., Bloh, K. M. and Kmiec, E. B. (2017). Insertional mutagenesis by CRISPR/Cas9 ribonucleoprotein gene editing in cells targeted for point mutation repair directed by short single-stranded DNA oligonucleotides. *PLoS ONE* **12**, e0169350.
- Rodgers, K. and McVey, M. (2016). Error-prone repair of DNA double-strand breaks. *J. Cell. Physiol.* **231**, 15-24.
- Salsman, J. and Deltre, G. (2017). Precision genome editing in the CRISPR era. *Biochem. Cell Biol.* **95**, 187-201.
- Shen, B., Zhang, X., Du, Y., Wang, J., Gong, J., Zhang, X., Tate, P. H., Li, H., Huang, X. and Zhang, W. (2013). Efficient knockin mouse generation by ssDNA oligonucleotides and zinc-finger nuclease assisted homologous recombination in zygotes. *PLoS ONE* **8**, e77696.
- Singh, P., Schimenti, J. C. and Bolcun-Filas, E. (2015). A mouse geneticist's practical guide to CRISPR applications. *Genetics* **199**, 1-15.
- Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y. E. and Zhang, J. (2016). RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun.* **7**, 10548.
- Srivastava, M., Nambiar, M., Sharma, S., Karki, S. S., Goldsmith, G., Hegde, M., Kumar, S., Pandey, M., Singh, R. K., Ray, P. et al. (2012). An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell* **151**, 1474-1487.
- Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. and Stahl, F. W. (1983). The double-strand-break repair model for recombination. *Cell* **33**, 25-35.
- Wang, K., Tang, X., Liu, Y., Xie, Z., Zou, X., Li, M., Yuan, H., Ouyang, H., Jiao, H. and Pang, D. (2016). Efficient generation of orthologous point mutations in pigs via CRISPR-assisted ssODN-mediated homology-directed repair. *Mol. Ther. Nucleic Acids* **5**, e396.
- Westerfield, M. (2000). *The Zebrafish Book: A Guide for the Laboratory use of Zebrafish (Danio rerio)*. Eugene: Univ. of Oregon Press.
- Won, M. and Dawid, I. B. (2017). PCR artifact in testing for homologous recombination in genomic editing in zebrafish. *PLoS ONE* **12**, e0172802.
- Würtele, H., Little, K. C. E. and Chartrand, P. (2003). Illegitimate DNA integration in mammalian cells. *Gene Ther.* **10**, 1791-1799.
- Xie, S.-L., Bian, W.-P., Wang, C., Junaid, M., Zou, J.-X. and Pei, D.-S. (2016). A novel technique based on in vitro oocyte injection to improve CRISPR/Cas9 gene editing in zebrafish. *Sci. Rep.* **6**, 34555.
- Yang, L., Guell, M., Byrne, S., Yang, J. L., De Los Angeles, A., Mali, P., Aach, J., Kim-Kiselak, C., Briggs, A. W., Rios, X. et al. (2013). Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* **41**, 9049-9061.
- Yoshimi, K., Kunihiro, Y., Kaneko, T., Nagahara, H., Voigt, B. and Mashimo, T. (2016). ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. *Nat. Commun.* **7**, 10431.
- Yu, C., Liu, Y., Ma, T., Liu, K., Xu, S., Zhang, Y., Liu, H., La Russa, M., Xie, M., Ding, S. et al. (2015). Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* **16**, 142-147.
- Zhang, Y., Zhang, Z. and Ge, W. (2018). An efficient platform for generating somatic point mutations with germline transmission in the zebrafish by CRISPR/Cas9-mediated gene editing. *J. Biol. Chem.* **293**, 6611-6622.
- Zorin, B., Hegemann, P. and Sizova, I. (2005). Nuclear-gene targeting by using single-stranded DNA avoids illegitimate DNA integration in *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **4**, 1264-1272.
- Zu, Y., Tong, X., Wang, Z., Liu, D., Pan, R., Li, Z., Hu, Y., Luo, Z., Huang, P., Wu, Q. et al. (2013). TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat. Methods* **10**, 329-331.

SUPPLEMENTARY FIGURES

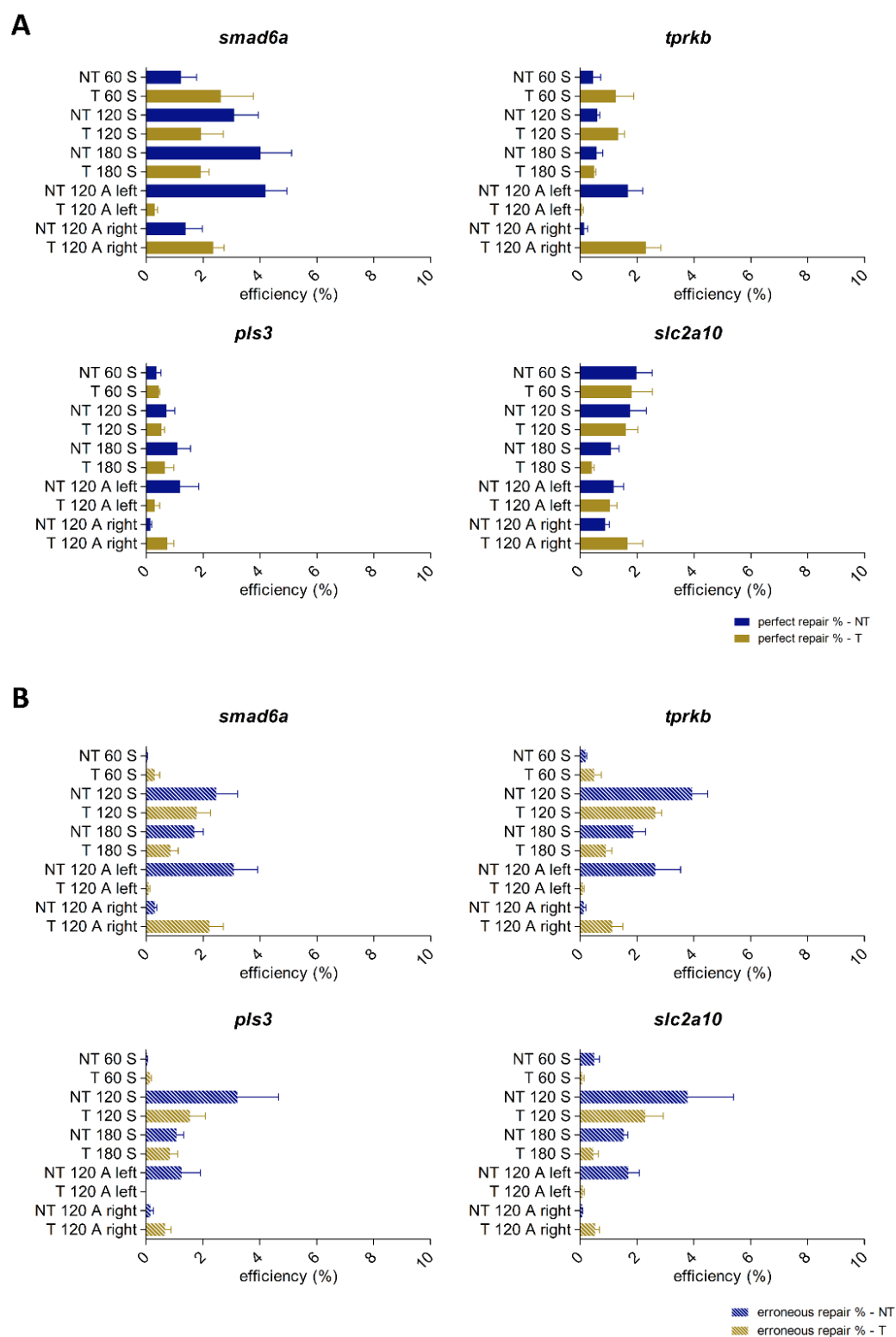


Fig. S1: Impact of repair template homology arm length, strand complementarity and symmetry on HDR efficiency. For each target site and each repair template type, average HDR efficiencies resulting from 5 repetitive experiments were plotted as perfect repair events (panel A) or erroneous repair events (panel B). NT: non-target, T: target, 60-120-180: total repair template length, S: symmetrical, A: asymmetrical. Error bars represent the SEM for 5 independent biological replicates each consisting of a pooled sample of 20 embryos.

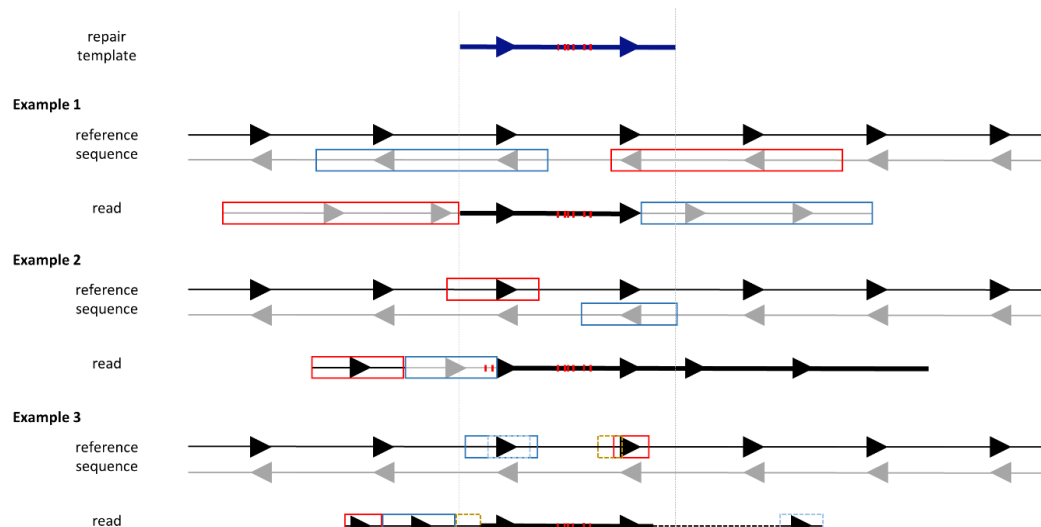
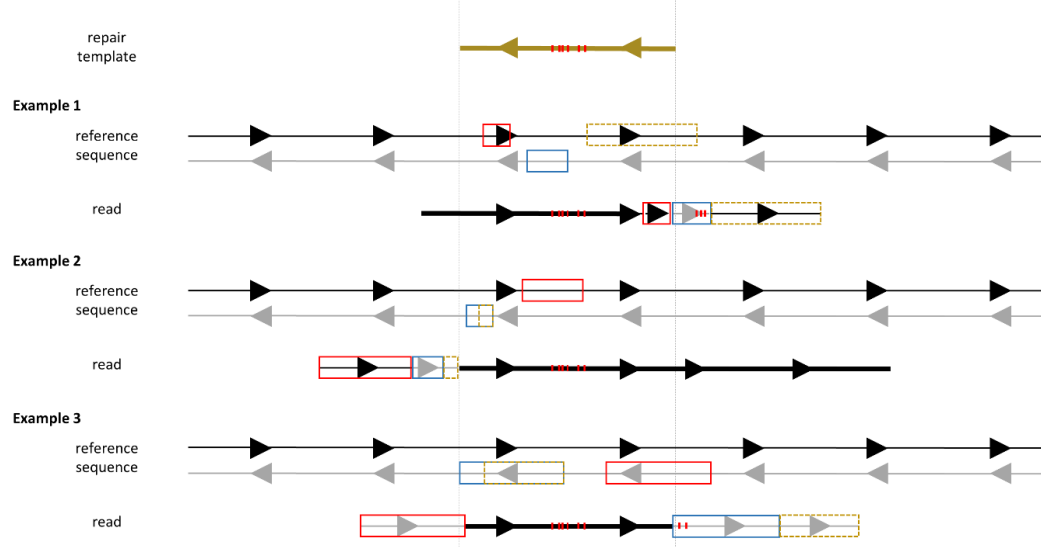
NT 120 S**T 120 S**

Fig. S2: Schematic representation of erroneous repair template integration events in the *smad6a* zebrafish gene with the NGS reads sequence as reference point. For both the 'NT 120 S' and 'T 120 S' repair templates, 3 examples of NGS reads are schematized to clarify the erroneous integration patterns that were encountered. For each 'Example' in the graph, both the reference sequence and the actual NGS read are depicted. The reference sequence consists of a black and a grey arrowed strand, representing the sense and antisense DNA strand, respectively. These strands serve as a reference frame used to estimate the size of the erroneous integration events and to indicate to which part of the reference sequence the integrated part of the repair template corresponds. The NGS read consists of both bold and thin arrowed segments. The bold segments correspond to the part of the read that actually mapped to the reference sequence during the NGS data analysis workflow. The thin lines correspond to the segments identified as non-matching (categorized as 'clipped off' by the CIGAR algorithm, see Fig. S4). Dashed segments represent an integration event, not corresponding to the sequence of the targeted gene. In both the repair template and the NGS read, precise base pair substitutions are indicated with red marks. Sequences corresponding to these examples are listed in Table S2.

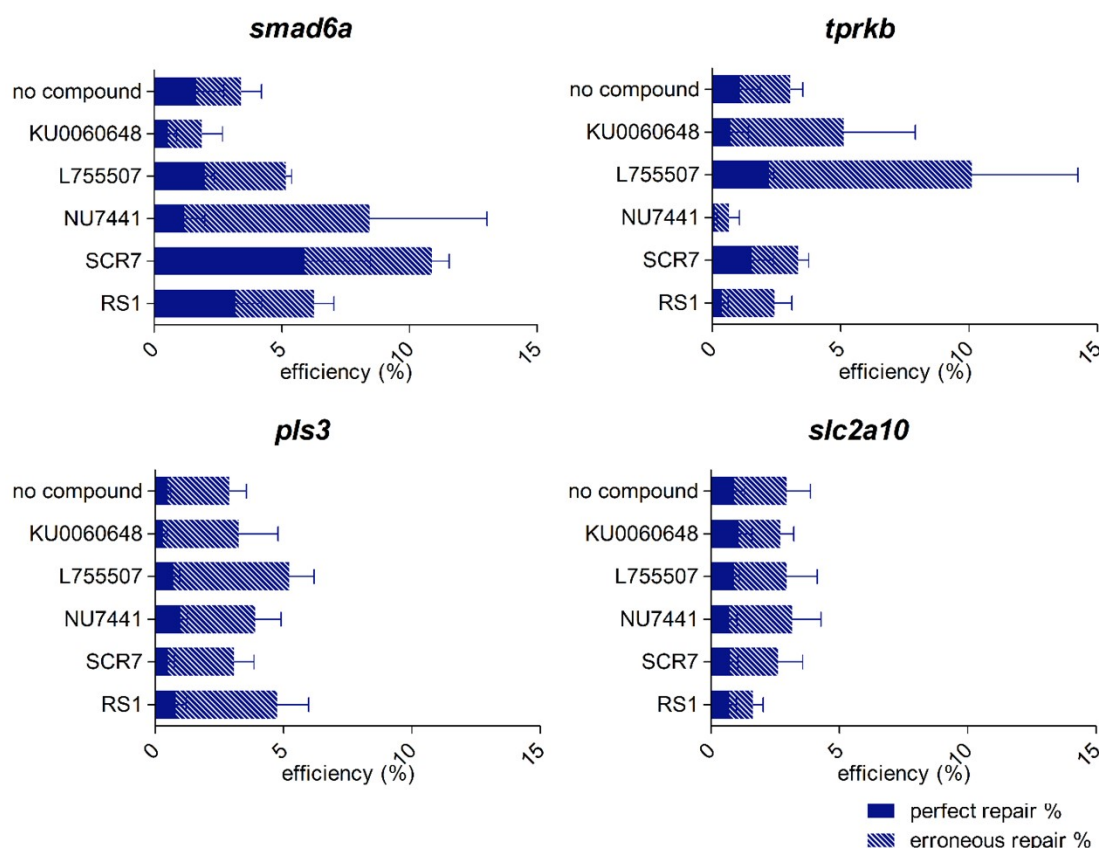


Fig. S3: Impact of chemical compound administration through incubation on HDR efficiency. Twenty embryos injected with mixes containing the NT 120 S repair template were incubated in screening medium supplemented with a chemical compound that either inhibit specific components of the NHEJ pathway, including SCR7, NU7441 and KU0060648 or that were shown to stimulate the HDR pathway, including RS1 and L755507. Three independent experiments were carried out and average total HDR rates are shown, split up in two categories: perfect repair % (plain bars) and erroneous repair % (dashed bars). Error bars represent the SEM for 3 independent biological replicates each consisting of a pooled sample of embryos. Indel rates depicted in this graph, are listed in Table S4. Statistical tests performed: independent samples t-test for normal distributed groups.

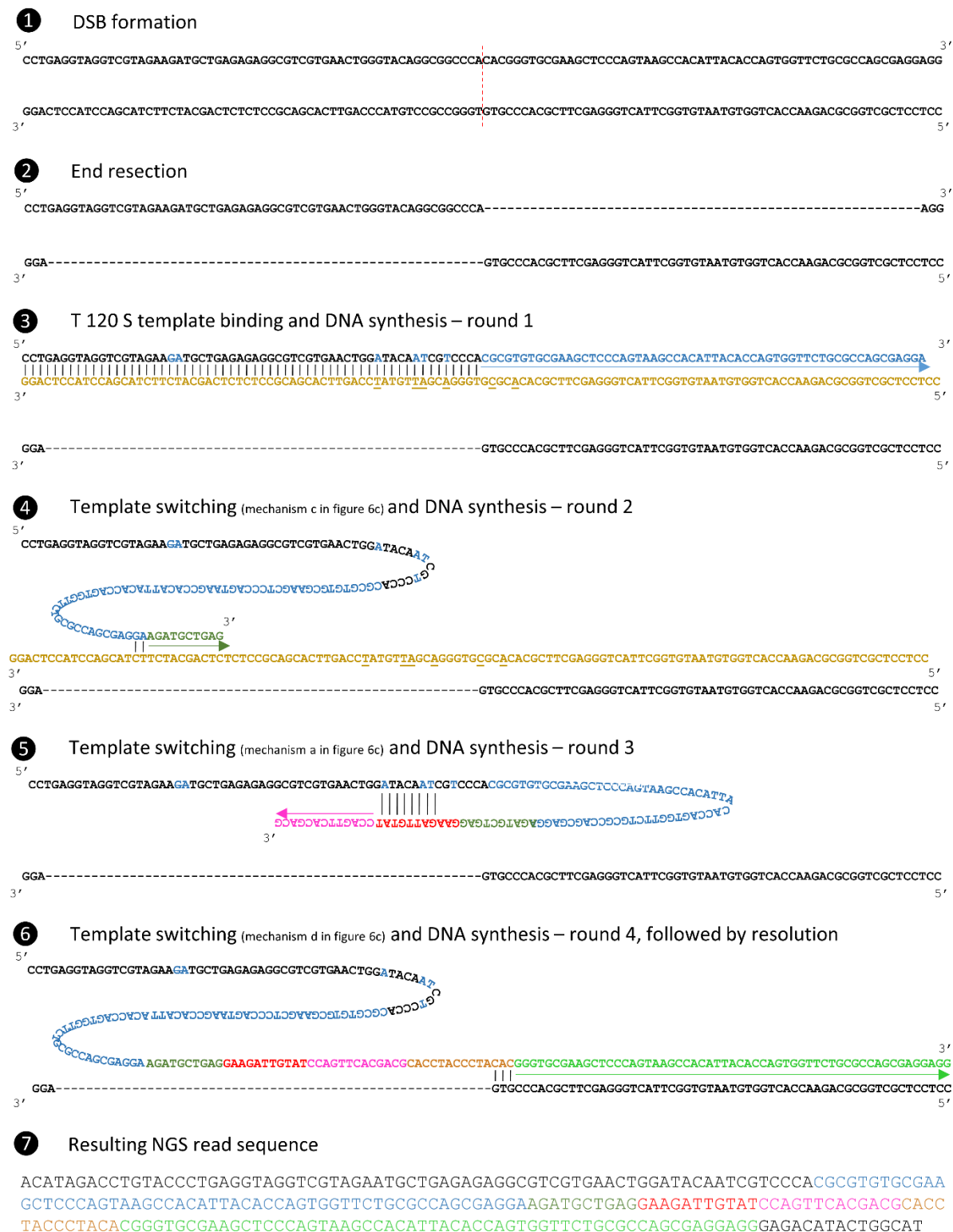


Fig. S4: Multiple template switching events occur during SDSA-dependent DSB repair. The origin of the complex mutational patterns encountered in this study is depicted, using an NGS read sequence, obtained from the analysis of the T 120 S repair template at the *smad6a* sgRNA target site. DSB formation (step 1) by Cas9 is followed by DNA end resection (step 2). Binding of the repair template (in yellow), is followed by DNA synthesis (blue arrow), leading to the lengthening of the 3' single-stranded DNA tail (blue bases) (step 3). Note the introduction of the intended precise base pair

substitutions, (underlined in the repair template). The complex dissociates and the 3' DNA tail, reanneals to the repair template, using a 2 bp microhomology pattern (step 4). This mechanism corresponds to 'mechanism c' depicted in Fig. 6C. From here DNA synthesis (green arrow) and further lengthening of the 3' single-stranded DNA tail (green bases) takes place. Red-colored bases are introduced via a non-identified template switching and DNA synthesis step. Complex dissociation and a reannealing step, using a 7 bp microhomology pattern, corresponding to 'mechanism a' depicted in Fig. 6C, leads to further lengthening of the 3' single-stranded DNA tail by DNA synthesis in the opposite direction (pink arrow and bases) (step 5). Again, brown-colored bases are introduced via a non-identified template switching and DNA synthesis step. Finally, reannealing according to 'mechanism d' depicted in Fig. 6C, using a 3 bp microhomology pattern, leads to a final DNA synthesis step (light green arrow and bases), followed by the resolution step (step 6). The resulting NGS read sequence, with color codes, is depicted at the bottom of the graph (step 7). The black-colored bases at the edges perfectly match the endogenous DNA sequence flanking the 120 depicted base pairs.

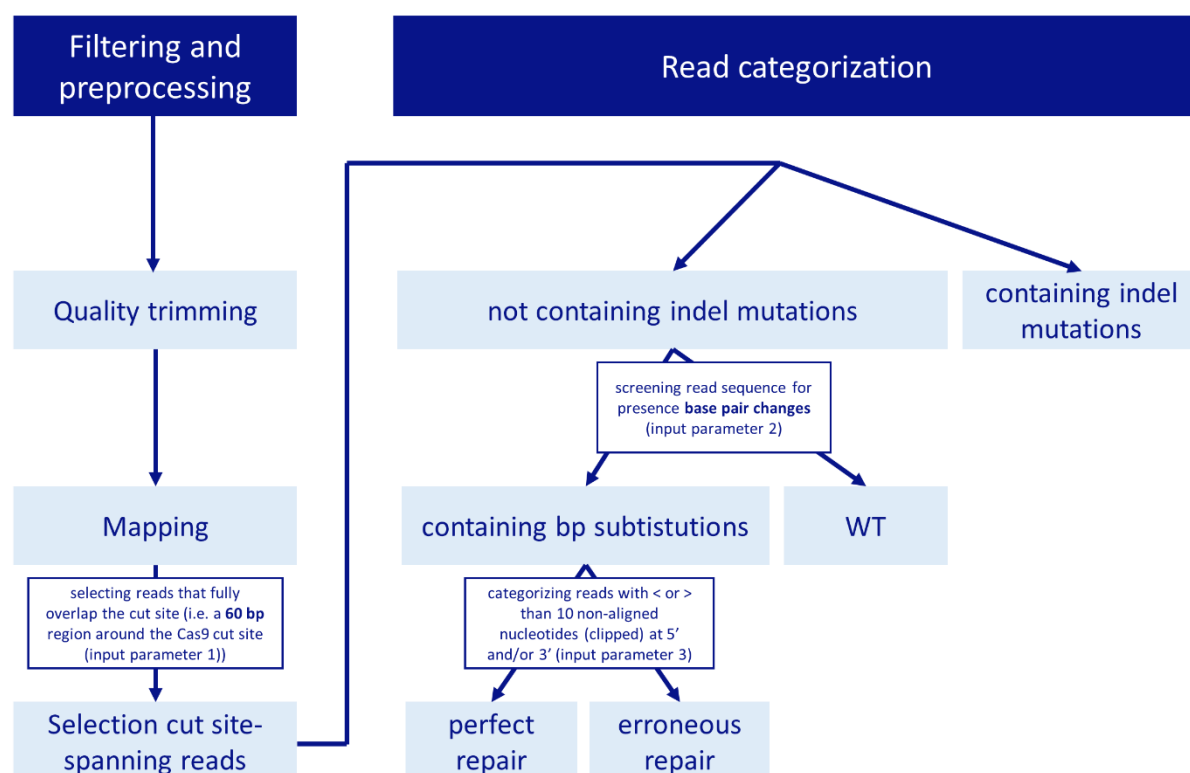


Fig. S5: Adapted BATCH-GE Next-Generation Sequencing data analysis pipeline for indel, perfect and erroneous repair rate calculation. Following quality trimming and read mapping, reads that fully overlap the cut site were divided into four categories (indel mutations, wild type, perfect repair and erroneous repair). Discrimination between perfect and erroneous repair is based on the number of nucleotides that are clipped from the alignment, i.e. ‘S’ or ‘H’ in the CIGAR specification for respectively soft and hard clipping. The BATCH-GE script needs specification of certain parameters. First, the precise screening region which is used in the last filtering step and the first categorizing step where the reads are divided into reads that do or do not contain any indels, can be modified (30 bp up and downstream of the theoretical cut site in our analyses – input parameter 1). Second, the BATCH-GE analysis method requires a repair template sequence as an input parameter for the HDR analysis. In our analyses, we supplied the repair template sequence fragment running from the first to the last base pair substitution (input parameter 2). A third input parameter is required as well, stating that an analysis discriminating between perfect and erroneous repair is desired (input parameter 3). Input parameters are also listed in Table S12.

SUPPLEMENTARY TABLES

Table S1: perfect and erroneous repair rates presented in Fig. 1B

sgRNA target	strand complementarity	length/symmetry	perfect repair (%)					erroneous repair (%)				
			1	2	3	4	5	1	2	3	4	5
<i>smad6a</i>	non-target	60 S	2,86	0,00	1,18	0,07	2,01	0,00	0,00	0,00	0,07	0,06
		120 S	5,97	1,54	2,12	4,04	1,78	0,16	2,56	1,98	2,69	4,89
		180 S	2,86	4,26	2,79	2,01	8,14	1,17	1,54	2,85	1,15	1,74
		120 A left	4,80	6,16	3,47	1,68	4,83	3,50	5,24	0,69	1,68	4,29
		120 A right	1,91	0,41	3,51	0,37	0,70	0,18	0,43	0,10	0,46	0,35
	target	60 S	7,05	0,90	0,94	1,52	2,68	0,07	0,00	0,94	0,46	0,00
		120 S	1,77	5,01	1,01	0,78	1,05	1,77	3,45	2,02	0,88	0,70
		180 S	2,25	2,43	2,47	0,90	1,48	1,29	0,29	1,51	0,13	1,02
		120 A left	0,29	0,65	0,21	0,31	0,00	0,00	0,16	0,21	0,00	0,00
		120 A right	1,70	1,38	3,27	3,15	2,28	1,61	3,55	2,87	2,32	0,78
<i>tprkb</i>	non-target	60 S	0,00	0,07	1,40	0,22	0,59	0,00	0,28	0,11	0,28	0,22
		120 S	0,39	0,78	0,68	0,55		3,92	5,49	3,40	2,94	
		180 S	1,42	0,28	0,48	0,44	0,25	2,85	0,83	2,71	2,07	0,87
		120 A left	1,68	1,09	0,75	3,64	1,24	1,68	0,54	1,50	4,55	4,97
		120 A right	0,00	0,06	0,65	0,00	0,00	0,00	0,25	0,11	0,33	0,00
	target	60 S	0,35	0,34	3,51	0,22	1,83	0,14	0,40	0,19	0,22	1,50
		120 S	1,48	1,89	0,85	1,08		2,63	2,43	3,28	2,28	
		180 S	0,63	0,32	0,61	0,41	0,48	0,76	1,62	0,79	0,93	0,48
		120 A left	0,26	0,00	0,00	0,00	0,00	0,20	0,24	0,00	0,00	0,00
		120 A right	2,68	1,04	4,20	1,85	1,74	2,01	0,00	1,87	1,11	0,63
<i>pls3</i>	non-target	60 S	0,28	0,91	0,36	0,26	0,00	0,00	0,08	0,09	0,06	0,00
		120 S	0,49	1,78	0,77	0,51	0,00	4,87	8,13	1,28	1,52	0,25
		180 S	1,18	2,82	0,26	0,80	0,46	0,39	1,69	1,55	1,20	0,46
		120 A left	0,00	1,81	3,50	0,33	0,32	0,00	1,81	3,50	0,00	0,96
		120 A right	0,13	0,33	0,14	0,08	0,00	0,00	0,57	0,00	0,16	0,00
	target	60 S	0,39	0,45	0,43	0,58	0,29	0,26	0,23	0,00	0,19	0,00
		120 S	0,39	0,52	0,34	0,37	1,00	0,20	0,77	1,37	3,18	2,23
		180 S	0,57	0,30	0,54	0,00	1,87	1,72	0,76	0,00	0,92	0,80
		120 A left	0,95	0,12	0,41	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<i>slc2a10</i>	non-target	120 A right	0,97	0,00	1,02	0,38	1,29	0,00	0,53	1,32	0,76	0,74
		60 S	1,37	1,78	0,51	3,68	2,57	0,81	0,11	0,07	0,99	0,51
		120 S	0,65	3,74	0,76	1,40	2,20	0,76	9,57	1,76	1,90	4,93
		180 S	0,88	1,00	1,42	1,89	0,23	1,12	1,55	1,78	1,89	1,32
		120 A left	1,29	2,26	0,65	0,20	1,45	1,93	3,11	1,30	1,01	1,09
	target	120 A right	0,57	0,62	1,26	1,15	0,78	0,05	0,06	0,17	0,06	0,00
		60 S	1,77	0,49	4,47	0,38	1,91	0,00	0,05	0,32	0,00	0,00
		120 S	0,46	1,46	0,99	2,85	2,23	1,84	1,46	3,81	0,62	3,74
		180 S	0,48	0,28	0,19	0,60	0,49	0,10	0,46	0,19	1,09	0,49
		120 A left	1,90	0,53	0,79	0,84	1,16	0,35	0,10	0,00	0,00	0,00
		120 A right	3,74	0,70	1,13	1,26	1,48	1,10	0,35	0,32	0,58	0,29

Table S2: sequence NGS reads depicted in Fig. 2 and Fig. S2

Repair template	Example	Read sequence
NT 120 S	1	TTAACATGTCGCTTCCACCTCTCTTTCAGATGCCAGTATGACTCCCTCCTCGCTGGCGCAGAACCACTGGTGTAATGTGGCTTACTGATGCTGAGAGAGGCGTCGTGAACTGGATACAATCGTCCCACGCGTGTGCGAAGCTCCAGTAAGCCACATTACACCGTGGTTCTGCGCCGCTGTACCCAGTTCACGACGCCTCTCTCAGCATCTTCTACGACCTACCTCAGGGTACAGGCCTATGTTTGGG
	2	TACCCTGAGGTAGGTCGTAGAGGATGCTGAGAGAGGCGTCGTGAACTGGGTACAGGCGGCCCTCCTCGCTGGGCGCAGAACCACTGGTGAATGTGGCTTACTGGGAGCTTCGCACACGCGTGGGAGGCGTCGTGAACTGGATACAATCGTCCCACGCGGGTGCGAAGCTCCAGTAAGCCACATTACACCACTGGTTCTGCGCCAGCGAGGAGGGAGACATACTGGCATCTGAAAGAGAGGTGGGAACG
	3	TTACACCACTGGTTCGTAGAAGATGCTGAGAGAGGCGTCGTGAACTACACATTACACATCGTAGAAGATGCTGAGAGAGGCGTCGTGAACTGGATACAATCGTCCCACGCGTGTGCGAAGCTCCAGTAAGCCGATTACACCACTGGTTCTGCGCCTCTGTGGTTCTAACCGTCGGATATCGGGAAGCTGAAGCTCCGACAGGTCTCACAGTGCGAAGGTCTCACTAATGTGAGAGAGGCGTCGTGAA
T 120 S	1	ACATAGACCTGTACCCTGAGGTAGGTCGTAGAATGCTGAGAGAGGCGTCGTGAACTGGATACAATCGTCCACGCGTGTGCGAAGCTCCAGTAAGCCACATTACACCACTGGTTCTGCGCCAGCGAGGAGGAGACATACTGGCAT
	2	CGTGAAGTGGGTACAGGCGGCCAGCATCTTCTACGAGCATCTTCTGAGGTAGGTCGTAGAAGATGCTGAGAGAGGCGTCGTGAACTGGATACAATCGTCCCACGCGTGTGCGAAGCTCCAGTAAGCCACATTACACCACTGGTTCTGCGCCAGCGAGGAGGAGACATACTAGCATCTGAAAGAGAGGTGGAAACGACATGTTAAGCATAAATTTAAAATAGAGAACACGAATCAAAATAAATAAAAT
	3	TATGTCTCCCTCCTCGCTGGCGCAGAACCACTGGTGAATGTGGCTTACTGAGGTAGGTCGTAGAAGATGCTGAGAGAGGCGTCGTGAACTGGATACAATCGTCCCACGCGTGTGCGAAGCTCCAGTAAGCCACATTACATTGTATCCAGTTCACGACGCCTCTCTCAGCATCTTCTACGACCTACCTCAGGTACCCAGTACCCAGTTCACGACGCCTCTCTCAGCATCTTCTACGAC

Table S3: perfect repair rates presented in Fig. 3

			perfect repair (%)																													
Experiment			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Position bp substitution			-14					-9					-8					-5					2					5				
smad6a	NT	60 S	0,00	0,00	0,08	0,00	0,45	0,00	0,00	0,00	0,07	0,65	0,00	0,00	0,00	0,07	0,65	0,00	0,00	1,18	0,07	0,71	2,86	0,00	1,18	0,07	2,01	2,86	0,00	1,26	0,07	1,82
		120 S	5,97	0,52	1,39	2,70	1,77	5,97	1,54	1,47	3,03	1,78	5,89	1,54	1,44	3,03	1,78	5,97	1,54	1,38	3,03	1,78	5,97	1,54	2,12	4,04	1,78	5,97	1,54	1,92	4,04	1,78
		180 S	0,93	2,87	1,69	1,43	8,14	0,99	3,43	2,79	1,72	8,14	0,79	3,43	2,92	1,65	8,14	0,94	4,12	2,86	1,80	8,14	2,86	4,26	2,79	2,02	8,14	2,65	4,26	2,72	1,73	6,94
		120 A left	4,28	5,85	3,48	1,40	4,29	4,75	5,86	3,48	1,68	4,29	4,75	6,01	3,48	1,68	4,29	4,75	6,16	3,48	1,68	4,29	4,81	6,16	3,48	1,68	4,82	4,53	5,99	2,09	0,84	1,87
	120 A right	0,49	0,02	3,29	0,05	0,18	0,85	0,02	3,31	0,05	0,53	0,80	0,02	3,27	0,05	0,53	0,76	0,12	3,24	0,06	0,53	1,92	0,41	3,51	0,37	0,70	1,43	0,25	3,34	0,26	0,53	
	T	60 S	0,15	0,21	0,00	0,61	0,19	6,52	0,28	0,32	0,91	1,72	6,52	0,28	0,32	0,91	1,72	6,52	0,28	0,62	0,91	1,72	7,04	0,90	0,94	1,51	2,68	6,67	0,55	0,62	1,36	2,68
		120 S	0,81	1,11	1,00	0,88	0,26	1,13	1,71	1,01	0,78	0,44	1,29	1,71	1,01	0,78	0,35	1,13	1,71	1,01	0,78	0,44	1,77	5,01	1,01	0,78	1,05	1,45	4,92	1,01	0,78	0,78
		180 S	0,24	0,44	1,51	0,26	0,78	0,80	1,33	2,36	0,45	1,05	0,80	1,25	2,36	0,45	1,05	0,88	1,33	2,36	0,48	1,02	2,25	2,43	2,47	0,90	1,47	1,85	1,98	2,47	0,55	1,41
120 A left		0,00	0,41	0,21	0,31	0,37	0,00	0,41	0,21	0,31	0,37	0,00	0,41	0,21	0,31	0,00	0,00	0,41	0,21	0,31	0,00	0,29	0,65	0,21	0,31	0,00	0,00	0,73	0,21	0,31	0,00	
120 A right	0,75	0,78	0,98	1,87	0,67	1,22	0,78	1,42	2,25	1,55	1,32	0,98	1,42	2,20	1,55	1,22	1,18	1,42	2,25	1,55	1,70	1,38	3,27	3,15	2,28	1,70	1,19	2,64	2,93	2,23		
Position bp substitution			-14					-8					2					5					8					11				
tprkb	NT	60 S	0,06	0,07	1,26	0,03	0,30	0,00	0,07	1,30	0,06	0,37	0,00	0,07	1,41	0,22	0,60	0,06	0,07	1,48	0,22	0,60	0,00	0,07	1,48	0,19	0,60	0,00	0,07	0,97	0,16	0,60
		120 S	0,47	0,76	0,79	0,37		0,39	0,78	0,69	0,55		0,39	0,78	0,68	0,55		0,39	0,78	0,79	0,37		0,39	0,76	0,69	0,37		0,39	0,53	0,68	0,18	
		180 S	0,28	0,42	0,32	0,37	0,13	0,28	0,28	0,48	0,29	0,25	1,43	0,28	0,48	0,44	0,25	1,71	0,42	0,55	0,44	0,38	1,43	0,28	0,48	0,44	0,25	1,41	0,28	0,32	0,44	0,38
		120 A left	0,84	0,54	0,75	2,73	0,00	1,05	1,08	0,75	3,64	1,24	1,68	1,08	0,75	3,64	1,24	1,68	1,08	0,75	3,64	1,24	1,68	1,08	0,75	3,64	1,24	0,84	0,55	0,75	0,90	1,22
		120 A right	0,20	0,00	0,11	0,00	0,00	0,00	0,00	0,11	0,05	0,00	0,00	0,06	0,65	0,00	0,00	0,00	0,00	0,76	0,00	0,00	0,00	0,00	0,65	0,00	0,00	0,00	0,00	0,65	0,00	0,00
	T	60 S	0,28	0,13	2,28	0,17	0,21	0,00	0,07	2,37	0,22	0,45	0,36	0,34	3,50	0,22	1,83	0,21	0,34	3,50	0,22	1,75	0,00	0,20	3,22	0,17	1,71	0,00	0,14	3,22	0,22	1,59
		120 S	0,52	0,54	0,48	0,12		0,62	0,54	0,61	0,23		1,48	1,89	0,85	1,08		1,37	1,62	0,98	0,96		1,48	1,62	0,98	0,96		1,48	1,36	0,98	0,84	
		180 S	0,38	0,54	0,53	0,41	0,35	0,38	0,00	0,44	0,10	0,14	0,63	0,32	0,62	0,42	0,49	0,88	0,32	0,78	0,31	0,62	0,38	0,32	0,62	0,31	0,55	0,51	0,32	0,53	0,52	0,49
120 A left	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,26	0,00	0,00	0,00	0,00	0,26	0,00	0,00	0,00	0,00	0,26	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
120 A right	0,67	0,00	2,15	0,37	0,48	0,67	0,00	2,15	0,55	0,78	2,68	1,04	4,20	1,85	1,74	2,68	1,04	4,05	1,85	1,74	2,67	1,04	4,20	1,65	1,74	1,35	1,04	4,20	1,85	1,74		
Position bp substitution			-15					-12					-3					1					10									
pls3	NT	60 S	0,14	0,49	0,09	0,06	0,17	0,78	0,74	0,27	0,58	0,87	0,28	0,99	0,36	0,32	0,17	0,28	0,91	0,36	0,39	0,17	0,21	0,58	0,27	0,13	0,00					
		120 S	0,49	1,34	0,51	0,76	0,25	0,56	1,56	0,51	0,76	0,51	0,63	1,66	0,77	0,51	0,25	0,49	1,77	0,77	0,51	0,00	0,21	0,85	0,77	0,26	0,51					
		180 S	0,79	1,69	0,51	0,40	0,46	0,79	2,82	0,51	1,19	0,92	1,19	1,69	0,51	0,79	0,46	1,19	2,82	0,26	0,80	0,46	0,79	1,69	0,26	0,40	0,00					
		120 A left	0,00	1,30	2,09	0,33	0,32	0,00	1,81	2,09	0,66	0,32	0,00	1,55	3,49	0,33	0,32	0,00	1,81	3,49	0,33	0,32	0,00	1,29	0,69	0,00	0,32					
		120 A right	0,00	0,19	0,29	0,04	0,00	0,54	0,29	0,21	0,40	0,00	0,27	0,34	0,14	0,08	0,00	0,27	0,34	0,14	0,08	0,00	0,13	0,19	0,14	0,12	0,00					
	T	60 S	0,00	0,11	0,00	0,10	0,00	0,13	0,68	0,86	0,49	0,96	0,26	0,23	0,29	0,58	0,07	0,39	0,45	0,43	0,58	0,29	0,00	0,11	0,29	0,68	0,07					
		120 S	0,20	0,52	0,00	0,00	0,44	0,59	0,52	0,00	0,00	0,66	0,00	0,52	0,35	0,18	1,01	0,39	0,52	0,35	0,38	1,01	0,20	0,52	0,35	0,38	0,89					
		180 S	0,19	0,15	0,00	0,00	0,27	0,95	0,90	0,27	0,00	0,80	0,77	0,30	0,27	0,00	0,54	0,57	0,30	0,54	0,00	1,87	0,77	0,46	0,27	0,00	0,27					
120 A left	0,00	0,00	0,20	0,00	0,00	0,00	0,47	0,41	0,00	0,00	0,00	0,12	0,41	0,00	0,00	0,95	0,12	0,41	0,00	0,00	0,00	0,23	0,00	0,00	0,00	0,00						
120 A right	0,32	0,00	0,15	0,19	0,18	0,65	1,60	0,58	0,75	1,09	1,30	0,00	1,16	0,38	0,92	0,97	0,00	1,03	0,38	1,29	0,32	0,53	0,73	0,38	0,92							
Position bp substitution			-15					-14					-13					-10					1					6				
slc2a10	NT	60 S	0,25	0,44	0,05	0,60	0,31	0,18	0,44	0,02	0,53	0,36	0,18	0,44	0,05	0,55	0,31	0,50	0,74	0,44	0,74	0,41	1,37	1,79	0,51	3,68	2,57	0,43	0,66	0,15	0,34	0,31
		120 S	0,49	1,24	0,26	0,60	1,16	0,42	1,24	0,26	0,50	1,22	0,47	1,22	0,33	0,50	1,16	0,85	1,49	0,33	1,20	1,39	0,65	3,74	0,76	1,39	2,21	0,29	1,04	0,17	0,29	0,46
		180 S	0,30	0,64	0,47	0,18	0,29	0,24	0,64	0,47	0,18	0,17	0,24	0,64	0,47	0,18	0,17	0,59	1,09	0,47	0,36	0,57	0,88	1,00	1,42	1,89	0,23	0,18	0,82	0,12	0,09	0,12
		120 A left	0,84	0,94	0,65	0,00	0,37	0,78	0,94	0,43	0,00	0,37	0,88	0,94	0,43	0,20	0,37	1,05	1,32	1,29	0,41	0,48	1,28	2,26	0,65	0,20	1,45	0,71	0,84	0,22	0,00	0,37
		120 A right	0,00	0,02	0,00	0,01	0,00	0,00	0,01	0,00	0,01	0,00	0,03	0,00	0,05	0,01	0,22	0,34	0,36	0,31	0,43	0,22	0,57	0,62	1,26	1,15	0,78	0,03	0,01	0,07	0,06	0,11
	T	60 S	0,00	0,05	1,97	0,13	0,94	0,00	0,00	1,97	0,17	0,90	0,00	0,05	1,94	0,26	0,94	0,25	0,39	2,22	0,60	1,32	1,77	0,49	4,47	0,38	1,91	0,04	0,15	1,26	0,34	1,28
		120 S	0,20	0,25	0,49	0,10	0,33	0,07	0,33	0,49	0,10	0,33	0,20	0,25	0,54	0,10	0,36	0,46	0,73	0,59	0,26	0,63	0,46	1,46	0,99	2,84	2,22	0,27	0,25	0,54	2,13	0,89
		180 S	0,00	0,09	0,06	0,44	0,33	0,00	0,05	0,13	0,38	0,29	0,00	0,05	0,13	0,38	0,29	0,19	0,32	0,70	1,09	0,66	0,48	0,28	0,19	0,60	0,49	0,10	0,19	0,13	0,27	0,29
120 A left	0,10	0,05	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,03	0,05	0,00	0,14	0,00	0,42	0,38	0,40	0,00	0,00	1,90	0,53	0,79	0,00	0,00	0,16	0,10	0,00	0,14	0,00		
120 A right	0,07	0,00	0,05	0,03	0,04	0,07	0,00	0,02	0,00	0,04	0,07	0,00	0,02	0,03	0,07	0,29	0,82	0,37	0,59	0,65	3,74	0,71	1,13	1,26	1,48	1,25	0,12	0,39	0,31	0,58		

Table S4: perfect and erroneous repair rates presented in Fig. 4 (compound injection)

gene	compound	perfect repair (%)					erroneous repair (%)				
		1	2	3	4	5	1	2	3	4	5
<i>smad6a</i>	no compound	5,97	1,54	2,12	4,04	1,78	0,16	2,56	1,98	2,69	4,89
	KU0060648	5,92	1,76	9,41	0,81	1,58	4,86	0,93	1,75	1,94	2,06
	L755507	5,34	1,48	2,59	3,48	5,96	1,93	6,49	2,65	2,45	2,55
	NU7441	3,38	5,71	1,19	2,67	1,37	0,75	1,27	4,07	2,03	0,68
	SCR7	2,16	1,07	0,89	2,65	0,68	3,72	1,16	0,78	4,76	0,68
	RS1	1,20	2,35	3,65	2,22	3,99	0,00	4,51	0,00	2,22	5,52
<i>tprkb</i>	no compound	0,39	0,78	0,68	0,55		3,92	5,49	3,40	2,94	
	KU0060648	0,87	2,35	0,79	0,58	2,34	5,81	7,06	6,15	0,00	2,34
	L755507	0,55	1,85	1,14	0,00	0,00	2,89	4,61	6,26	0,00	2,94
	NU7441	0,00	0,00	1,29	0,00	0,00	2,15	3,25	3,38	0,68	1,64
	SCR7	0,00	0,24	0,50	0,00	0,00	2,49	0,47	1,00	11,25	8,51
	RS1	0,00	1,95	0,00	0,00	1,89	0,00	3,25	1,69	2,61	1,89
<i>pls3</i>	no compound	0,49	1,78	0,77	0,51	0,00	4,87	8,13	1,28	1,52	0,25
	KU0060648	1,75	0,66	1,42	0,58	1,23	5,24	7,23	8,78	3,69	3,70
	L755507	1,40	0,55	0,54	1,23	1,54	4,72	4,98	4,70	2,87	1,54
	NU7441	1,16	2,52	0,88	0,37	0,41	1,45	9,88	11,82	1,12	2,06
	SCR7	1,73	1,46	9,43	1,28	1,72	10,64	5,33	5,66	4,70	0,00
	RS1	2,42	0,44	2,16	0,00	0,00	2,42	3,95	4,58	6,28	0,76
<i>slc2a10</i>	no compound	0,65	3,74	0,76	1,40	2,20	0,76	9,57	1,76	1,90	4,93
	KU0060648	0,79	2,40	0,99	1,20	3,06	1,64	2,99	1,54	3,23	7,87
	L755507	0,09	2,02	0,42	0,97	4,02	0,89	2,26	0,60	1,74	2,68
	NU7441	0,00	0,00	0,90	1,23	2,07	0,00	1,21	3,01	4,53	2,53
	SCR7	1,03	1,07	0,80	0,76	0,00	3,55	3,13	1,84	0,95	1,83
	RS1	0,81	3,62	1,62	1,44	4,06	2,03	1,88	1,98	3,31	6,35

Table S5: perfect and erroneous repair rates presented in Supplementary Fig. S3 (compound incubation)

gene	compound	perfect repair (%)			erroneous repair (%)		
		1	2	3	1	2	3
<i>smad6a</i>	no compound	1,33	3,64	0,00	2,80	2,27	0,17
	KU0060648	1,19	0,00	0,39	2,83	0,00	1,17
	L755507	1,33	1,97	2,63	2,87	3,03	3,64
	NU7441	0,88	2,68	0,00	4,27	1,17	16,30
	SCR7	2,47	4,30	10,93	5,07	3,75	6,11
	RS1	1,45	4,98	3,14	4,55	1,87	2,79
<i>tprkb</i>	no compound	2,63	0,61	0,00	2,63	1,02	2,26
	KU0060648	0,00	2,12	0,00	1,53	1,69	10,00
	L755507	1,89	2,58	2,15	16,04	2,58	5,12
	NU7441	0,30	0,00	0,00	0,30	0,00	1,36
	SCR7	0,00	1,74	2,89	1,52	2,62	1,27
	RS1	0,87	0,23	0,00	0,87	3,17	2,17
<i>pls3</i>	no compound	0,67	0,26	0,54	3,19	2,93	1,08
	KU0060648	0,56	0,35	0,00	1,83	5,98	1,04
	L755507	1,12	0,25	0,75	4,31	2,97	6,29
	NU7441	1,08	0,47	1,39	2,94	1,17	4,65
	SCR7	0,88	0,00	0,60	4,09	1,42	2,19
	RS1	1,43	0,95	0,00	2,04	6,20	3,68
<i>slc2a10</i>	no compound	1,62	0,25	0,87	3,62	0,39	2,08
	KU0060648	0,18	1,13	1,96	0,57	2,04	2,24
	L755507	1,29	0,57	0,82	4,42	0,64	1,12
	NU7441	0,94	0,08	1,09	3,11	0,23	4,00
	SCR7	1,07	0,14	1,04	3,38	0,07	2,13
	RS1	0,16	0,95	1,03	0,10	1,28	1,37

Table S6: perfect and erroneous repair rates presented in Fig. 5

gene	fish #	perfect repair (%)	erroneous repair (%)
<i>smad6a</i>	1	13,75	0,13
	2	22,97	0,00
	3	8,51	0,00
	4	0,26	0,26
	5	0,70	22,06
	6	0,48	0,00
	7	10,65	21,82
	8	0,14	0,14
<i>slc2a10</i>	1	0,00	1,82
	2	10,33	6,04
	3	0,15	0,00
	4	0,04	0,00
	5	0,05	0,00
	6	0,19	0,00
	7	0,27	0,40
	8	0,17	0,08

Table S7: gBlock designs for *tprkb*, *pls3*, *slc2a10* and *smad6a*

gene	exon	strand	protospacer + <u>PAM</u>	double-stranded DNA (gBlock) sequence (5'-3')
<i>tprkb</i>	2	+	GGATGCATTCCAGATCCTGGT <u>TGG</u>	CCGCTAGCTAATACGACTCACTATAGGATGCATTCCAGATCCT GGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCAGTCGGTGCTTTT
<i>pls3</i>	7	+	GGCTGTCACGTGGTCAACAT <u>TGG</u>	CCGCTAGCTAATACGACTCACTATAGGCTGTCACGTGGTCAA CATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGT CCGTTATCAACTTGAAAAAGTGGCACCAGTCGGTGCTTTT
<i>slc2a10</i>	2	+	AAAGCAAAGATAACATGCGG <u>GAGG</u>	CCGCTAGCTAATACGACTCACTATAGGAGCAAAGATAACATG CGGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGT CCGTTATCAACTTGAAAAAGTGGCACCAGTCGGTGCTTTT
<i>smad6a</i>	4	-	GGGTACAGGCGGCCACAC <u>GCGG</u>	ACCGCTAGCTAATACGACTCACTATAGGGTACAGGCGGCCCA CACGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGT CCGTTATCAACTTGAAAAAGTGGCACCAGTCGGTGCTTTT

Table S8: single-stranded oligodeoxynucleotide (ssODN) designs for *tprkb*, *pls3*, *slc2a10* and *smad6a*

gene	orientation	length (nt)	ssODN sequence (5'-3')
<i>tprkb</i>	non-target	60	TAATAAAGGTTGTGGACGCATTTAGATCCTCGTAGCCACCAATAAAGCAGTTCACCTTC
		120	TGTACGTATATTACCTTTTTATTTTTTTTAAATAAAGGTTGTGGACGCATTTAGATCCTCGTAGCCACC AATAAAGCAGTTCACCTTCATAAAATCGGGAAAATGAAACCAGAAGCC
		180	AAATAATAGTCAAAATATATTCTAAAAACATGTACGTATATTACCTTTTTATTTTTTTTAAATAAAGGTT GTGGACGCATTTAGATCCTCGTAGCCACCAATAAAGCAGTTCACCTTCATAAAATCGGGAAAATGAAA ACCAGAAGCCTTTACTCTGAAATCATTTTCAATCTTTCAC
		left	AAATAATAGTCAAAATATATTCTAAAAACATGTACGTATATTACCTTTTTATTTTTTTTAAATAAAGGTT GTGGACGCATTTAGATCCTCGTAGCCACCAATAAAGCAGTTCACCTTC
		right	TAATAAAGGTTGTGGACGCATTTAGATCCTCGTAGCCACCAATAAAGCAGTTCACCTTCATAAAATCG GGAAAATGAAACCAGAAGCCTTTACTCTGAAATCATTTTCAATCTTTCAC
		60	GAAGGTGAACCTGCTTTATTGGTGGCTACGAGGATCTGAAATGCGTCCACAACCTTTATTA
	target	120	GGCTTCTGGTTTTCTTTTCCCGATTTTATGAAGGTGAACCTGCTTTATTGGTGGCTACGAGGATCTGAAA TGCGTCCACAACCTTTATTAATAAAAAATAAAGGTAAATATACGTACA
		180	GTGAAAGATTGAAATGATTTAGAGTAAAGGCTTCTGGTTTTCTTTTCCCGATTTTATGAAGGTGAA CTGCTTTATTGGTGGCTACGAGGATCTGAAATGCGTCCACAACCTTTATTAATAAAAAATAAAGGTAA AATATACGTACATGTTTTAGAATATATTTGACTATTATTT
		left	GAAGGTGAACCTGCTTTATTGGTGGCTACGAGGATCTGAAATGCGTCCACAACCTTTATTAATAAAAAAT AAAAAGGTAATATACGTACATGTTTTAGAATATATTTGACTATTATTT
		right	GTGAAAGATTGAAATGATTTAGAGTAAAGGCTTCTGGTTTTCTTTTCCCGATTTTATGAAGGTGAA CTGCTTTATTGGTGGCTACGAGGATCTGAAATGCGTCCACAACCTTTATTA
		60	AGCCTCTGCCATTGGGTGCCACGTGGTGAATATTGAGCATTAGACCTTCGAGAAGGGAA
<i>pls3</i>	non-target	120	GCAGGAGAATCTTAACCTGGCTCTCAATTCAGCCTCTGCCATTGGGTGCCACGTGGTGAATATTGGAGC ATTAGACCTTCGAGAAGGGAAACCTCATCTGGTTTTAGGGCTTCTGTGGCA
		180	TGCTATATTGAAAAACATTCTTTCCATCCTGCAGGAGAATCTTAACCTGGCTCTCAATTCAGCCTCTGCC ATTGGGTGCCACGTGGTGAATATTGGAGCATTAGACCTTCGAGAAGGGAAACCTCATCTGGTTTTAGG GCTTCTGTGGCAGATCATCAAGATTGGTCTGTTGCTGACAT
		left	TGCTATATTGAAAAACATTCTTTCCATCCTGCAGGAGAATCTTAACCTGGCTCTCAATTCAGCCTCTGCC ATTGGGTGCCACGTGGTGAATATTGGAGCATTAGACCTTCGAGAAGGGAA
		right	AGCCTCTGCCATTGGGTGCCACGTGGTGAATATTGGAGCATTAGACCTTCGAGAAGGGAAACCTCATC TGGTTTTAGGGCTTCTGTGGCAGATCATCAAGATTGGTCTGTTGCTGACAT
	target	60	TTCCCTTCTCGAAGGTCTAATGCTCCAATATTCACCAGTGGCACCCAATGGCAGAGGCT
		120	TGCCACAGAAGCCCTAAAACAGATGAGGTTCCCTTCTCGAAGGTCTAATGCTCCAATATTCACCACG TGGCACCCAATGGCAGAGGCTGAATTGAGAGCCAAGTTAAGATTCTCTCTGC
		180	ATGTCAGAAAACAGACCAATCTTGATGATCTGCCACAGAAGCCCTAAAACAGATGAGGTTCCCTTCT CGAAGGTCTAATGCTCCAATATTCACCAGTGGCACCCAATGGCAGAGGCTGAATTGAGAGCCAAGTT AAGATTCTCTGAGGATGAAAGAATGTTTTTCAATATAGCA
		left	TTCCCTTCTCGAAGGTCTAATGCTCCAATATTCACCAGTGGCACCCAATGGCAGAGGCTGAATTGAGA GCCAAGTTAAGATTCTCTGAGGATGAAAGAATGTTTTTCAATATAGCA
		right	ATGTCAGAAAACAGACCAATCTTGATGATCTGCCACAGAAGCCCTAAAACAGATGAGGTTCCCTTCT CGAAGGTCTAATGCTCCAATATTCACCAGTGGCACCCAATGGCAGAGGCT
<i>slc2a10</i>	non-target	60	TCTGACCTTTTTAAATCAAAGGATAACATGAGGAGAAGAACGGTGATTGGTGTGGGTTG
		120	CAGCATCAGCAAAGTGAGAAATATGGTGTCTGACCTTTTTAAATCAAAGGATAACATGAGGAGAAG AACGGTGATTGGTGTGGGTTGGTGTGAGTCAGCAGTTCACAGGTCAACCA
		180	AGATTAACAGAGGAAACGGAAACATCAATCAGCATCAGCAAAGTGAGAAATATGGTGTCTGACCT TTTTAAATCAAAGGATAACATGAGGAGAAGAACGGTGATTGGTGTGGGTTGGTGTGAGTCAGCAG TTCACAGGTCAACCAATGTGCTCTTTACGCCTCAACTATACTT
		left	AGATTAACAGAGGAAACGGAAACATCAATCAGCATCAGCAAAGTGAGAAATATGGTGTCTGACCT TTTTAAATCAAAGGATAACATGAGGAGAAGAACGGTGATTGGTGTGGGTTG
		right	TCTGACCTTTTTAAATCAAAGGATAACATGAGGAGAAGAACGGTGATTGGTGTGGGTTGGTGTGAG TCAGCAGTTCACAGGTCAACCAATGTGCTCTTTACGCCTCAACTATACTT
	target	60	CAACCAACACCAATCACCCTTCTCTCCTCATGTTATCCTTTGATTAAAAAGGTCAGA
		120	TGGTTGACCTGTGAACCTGCTGACTCAGCACCAACCAACCAATCACCCTTCTCTCCTCATGTTATCC TTTGATTAAAAAGGTCAGAAACCATATTTCTCACTTTGCTGATGCTG
		180	AAGTATAGTTGAGGCGTAAAGAGCACATTTGGTTGACCTGTGAACCTGCTGACTCAGCACCAACCA ACCAATCACCCTTCTCTCCTCATGTTATCCTTTGATTAAAAAGGTCAGAAACCATATTTCTCACTT TGCTGATGCTGATTGATGTTCCGTTTCCTCTGTTAATCT

smad6a		left	CAACCCAACACCAATCACCGTTCTTCTCCTCATGTTATCCTTTGATTTAAAAAGGTCAGAAACACCATAT TTCTCACTTTGCTGATGCTGATTGATGTTCCGTTTCCTCTGTAAATCT
		right	AAGTATAGTTGAGGCGTAAAAGAGCACATTTGTTGACCTGTGAAGTCTGACTCAGCACCAACCCAA CACCAATCACCGTTCTTCTCCTCATGTTATCCTTTGATTTAAAAAGGTCAGA
	non-target	60	AGGCGTCGTGAACTGGATACAATCGTCCACGCGTGTGCGAAGCTCCAGTAAGCCACAT
		120	CCTGAGGTAGGTCTAGAGAAGTGTGAGAGAGGCGTCTGAACTGGATACAATCGTCCACGCGTGT GCGAAGCTCCAGTAAGCCACATTACACCAGTGGTCTGCGCCAGCGAGGAGG
		180	GGGGCAGCAGGCCCAAACATAGGCCTGTACCCTGAGGTAGGTCTGAGAAGATGCTGAGAGAGGCGTC GTGAACTGGATACAATCGTCCACGCGTGTGCGAAGCTCCAGTAAGCCACATTACACCAGTGGTCT GCGCCAGCGAGGAGGGAGACATACTGGCATCTGAAAGAGAGGTGG
		left	GGGGCAGCAGGCCCAAACATAGGCCTGTACCCTGAGGTAGGTCTGAGAAGATGCTGAGAGAGGCGTC GTGAACTGGATACAATCGTCCACGCGTGTGCGAAGCTCCAGTAAGCCACAT
		right	AGGCGTCGTGAACTGGATACAATCGTCCACGCGTGTGCGAAGCTCCAGTAAGCCACATTACACCAG TGGTTCTGCGCCAGCGAGGAGGGAGACATACTGGCATCTGAAAGAGAGGTGG
	target	60	ATGTGGCTTACTGGGAGCTTCGCACACGCGTGGGACGATTGTATCCAGTTCACGACGCCT
		120	CCTCTCGTGGCGCAGAACCACTGGTGTAAATGTGGCTTACTGGGAGCTTCGCACACGCGTGGGACGA TTGTATCCAGTTCACGACGCCTCTCTCAGCATCTTCTACGACCTACCTCAGG
		180	CCACCTCTCTTTCAGATGCCAGTATGTCTCCCTCCTCGTGGCGCAGAACCACTGGTGTAAATGTGGCTTA CTGGGAGCTTCGCACACGCGTGGGACGATTGTATCCAGTTCACGACGCCTCTCTCAGCATCTTCTACGA CCTACCTCAGGGTACAGGCCTATGTTTGGGCCTGCTGCCCC
		left	ATGTGGCTTACTGGGAGCTTCGCACACGCGTGGGACGATTGTATCCAGTTCACGACGCCTCTCTCAGCA TCTTCTACGACCTACCTCAGGGTACAGGCCTATGTTTGGGCCTGCTGCCCC
		right	CCACCTCTCTTTCAGATGCCAGTATGTCTCCCTCCTCGTGGCGCAGAACCACTGGTGTAAATGTGGCTTA CTGGGAGCTTCGCACACGCGTGGGACGATTGTATCCAGTTCACGACGCCT

Table S9: injection mix compositions

Component	Injection mixes					
	No compound	SCR7	NU7441	KU0060648	RS1	L755507
Cas9 (1000 ng/ μ l)	0,7 μ l	0,7 μ l	0,7 μ l	0,7 μ l	0,7 μ l	0,7 μ l
sgRNA (100 ng/ μ l)	0,7 μ l	0,7 μ l	0,7 μ l	0,7 μ l	0,7 μ l	0,7 μ l
phenol red (2%)	1 μ l	1 μ l	0,8 μ l	0,8 μ l	1 μ l	0,8 μ l
ssODN (10 μ M)	0,8 μ l	0,8 μ l	0,8 μ l	0,8 μ l	0,8 μ l	0,8 μ l
compound	0 μ l	0,4 μ l (10 mM)	1 μ l (2 mM)	1 μ l (0,5 mM)	0,2 μ l (20 mM)	1 μ l (2,5 mM)
nuclease-free water	0,8 μ l	0,4 μ l	0 μ l	0 μ l	0,6 μ l	0 μ l

Table S10: Selected concentrations compound incubation

Compound	Administered concentration (μ M)
SCR7	5
NU7441	2
KU0060648	2
RS1	100
L755507	2

Table S11: PCR primers for *tprkb*, *pls3*, *slc2a10* and *smad6a*

gene	primer sequence (5'-3')	
<i>tprkb</i>	forward	TCACCCAAGCACATAACA
	reverse	CGCCAACATAGGGAGAACAT
<i>pls3</i>	forward	ACCCTGTTTGTTCCTCGTA
	reverse	TTGTCATTACCATGCCCAAG
<i>slc2a10</i>	forward	TCACGTTGGCATCTTGATA
	reverse	AAGTGCATCGTTTCGTTGTG
<i>smad6a</i>	forward	GCCAGGACTCAAAACAAGC
	reverse	CACACTGTTCTGGGTCGTATG

Table S12: input parameters BATCH-GE

gene	Input parameter 1			Input parameter 2	Input parameter 3
	Chromosome	Start position	End position	Repair template sequence	Filter integration (yes/no)?
<i>tprkb</i>	chr4	28383726	28383785	(C)GCATT(T)CAGATCCT[C]GT(A)GC(C)AC(C)	yes
<i>pls3</i>	chr14	13159801	13159860	(G)TG(C)CACGTGGT(G)AA[T]ATTGGAGC(A)	yes
<i>slc2a10</i>	chr11	2348500	2348559	(T)(C)(A)AA(G)GATAACATG[A]GGAG(A)	yes
<i>smad6a</i>	chr7	33912937	33912996	(A)TACA(A)(T)CG(T)CCCAC[G]CG(T)G	yes

FIRST PERSON

First person – Annekatrien Boel

First Person is a series of interviews with the first authors of a selection of papers published in Disease Models & Mechanisms, helping early-career researchers promote themselves alongside their papers. Annekatrien Boel is first author on 'CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments', published in DMM. Annekatrien is a PhD student in the lab of Paul Coucke at Ghent University, Belgium, investigating disease modeling using CRISPR/Cas9 genome editing, with a specific interest in the generation of knock-in models harboring disease-relevant point mutations.

How would you explain the main findings of your paper to non-scientific family and friends?

The study of genetic disorders relies heavily on the use of appropriate animal models of disease. Until recently, the generation of such models, harboring disease-causing mutations in the gene of interest, was quite challenging. Newly emerged genome editing techniques, of which the CRISPR/Cas9 system is most notable, have facilitated disease model generation; however, technical challenges still hamper large-scale and efficient disease model development. Many disorders are caused by specific point mutations, which are alterations in one single nucleotide pair in the DNA. The introduction of such mutations requires the administration of the components of the CRISPR/Cas9 system, alongside a repair template, which is a piece of DNA, containing the base pair alteration the researchers aim to introduce, in a process called 'homology-directed repair' (HDR). To optimize this procedure, we employed the zebrafish, which is an increasingly attractive animal to model and study human disease. Two main findings can be extracted from our study. Firstly, we observed that, in addition to the introduction of the base pair alteration of interest, complex mutational patterns of repair template fragments were introduced in the genome. Secondly, the extent to which this erroneous repair was present, was dependent on the composition of the repair template that was used.

What are the potential implications of these results for your field of research?

The finding that, and the extent to which, repair template fragments were introduced into the genome during HDR was unexpected. The technique of administering the CRISPR/Cas9 complex along with a repair template is applied in many model organisms, as well as in human cells, and might be eventually used in therapeutic applications. Therefore, it is important to recognize the manifestation of erroneous repair during HDR, which was until now generally considered to be error-free, and to apply the appropriate techniques to analyze the efficiency and correctness of editing.

Annekatrien Boel's contact details: Center for Medical Genetics, Ghent University, C. Heymanslaan 10, 9000 Ghent, Belgium.
E-mail: Annekatrien.Boel@UGent.be



Annekatrien Boel

What are the main advantages and drawbacks of the model system you have used as it relates to the disease you are investigating?

"These characteristics make zebrafish the ideal *in vivo* disease model to conduct large-scale genome editing studies."

The main strength of zebrafish, in comparison to pre-eminently used model systems, such as rodents, is its high-throughput capability, displaying a high fecundity, *ex vivo* fertilization, and a rapid and *ex utero* development. These characteristics make zebrafish the ideal *in vivo* disease model to conduct large-scale genome editing studies. One of its main advantages can also turn into a disadvantage: a rapid development means that zebrafish undergo fast consecutive cell division cycles, resulting in highly mosaic F0 animals harboring a multitude of different mutations.

What has surprised you the most while conducting your research?

In one specific experiment of our study, we investigated the influence of five chemical compounds previously shown to improve HDR rates in cellular systems. Surprisingly, at four different target sites in the zebrafish genome, none of the administered compounds had a notable influence on HDR rates. While this outcome was rather disappointing, we believed it was still useful to include these results in our manuscript. Publishing experimental set-ups that do not work might save other research groups precious time and effort.

“Publishing experimental set-ups that do not work might save other research groups precious time and effort.”**Describe what you think is the most significant challenge impacting your research at this time and how will this be addressed over the next 10 years?**

While the emergence of CRISPR/Cas9 genome editing is one of the most exciting developments of recent years in the field of genetics,

the way to its implementation in medicine is still long. Findings like ours absolutely do not undermine the power of this technique. They do, however, contribute to the establishment of CRISPR/Cas9 as a safe technique by pinpointing the challenges that come along with it. In coming years, researchers will continue to do so and each possible risk regarding specificity or accuracy should be addressed appropriately in future applications. The ever-decreasing cost of next-generation sequencing techniques will be absolutely beneficial in this regard.

What's next for you?

After completion of my PhD, I will start work as a postdoctoral fellow, supporting different research groups that aim to implement the CRISPR/Cas9 genome editing technique in their research projects.

Reference

Boel, A., De Saffel, H., Steyaert, W., Callewaert, B., De Paepe, A., Coucke, P. J. and Willaert, A. (2018). CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments. *Dis. Model. Mech.* 11: dmm035352.